IDENTIFICATION OF EPIGENETICALLY DYSREGULATED GENES IN TUMOURS THAT METASTASISE TO THE BRAIN

A thesis presented in fulfilment of the requirements for the degree of Doctor of Philosophy (Ph.D.) of the University of Wolverhampton

Rajendra Prasad Pangeni
(Student No: 0924555)

Research Institute in Healthcare Sciences (RIHS)
Faculty of Science and Engineering (FoSE)
University of Wolverhampton
July 2015

Declaration

I hereby declare that this project thesis is written completely by myself and the work submitted in this dissertation is the result of my own investigation except where otherwise stated. The information used from other sources has been appropriately referenced and acknowledged.

Signature:

Date:

Name: Rajendra Prasad Pangeni

Student No: 0924555

Course: Doctor of Philosophy

Project title: **Identification of epigenetically dysregulated genes in tumours that metastasise to the brain**

Director of Studies: Dr. Mark R Morris

Acknowledgement

Foremost, I would like to take this opportunity to express my earnest gratitude to my supervisor Dr. Mark Morris for his wonderful support, guidance and rigorous training. I am immensely impressed by his wisdom and innovative approaches. It was a rare coincidence to find a good human and a scientist like him in a person. It was always an enormous joy and a great experience to be part of his research.

I am grateful to Dr. Tracy Warr, my second supervisor, for her support and valuable suggestions she extended during this period. My sincere thanks go to Prof. John Darling for providing me with funding to attend conferences including Society of Neuro-Oncology (SNO) conference in Miami, USA last year. I am thankful to Prof. Weiguang Wang for providing me with breast cancer cell lines and letting me use his laboratory reagents and materials whenever I needed. I am thankful to Dr. Angel Armesilla for his valuable suggestions during the project.

I have a deep appreciation for my wife, Sheela for her all time untiring support during the project. Her contant love and support always helped me to complete the difficult task of writing this thesis. I am greatly indebted to my parents, family members, and relatives for their support and contributions, especially my mother, who took good care of my daughter for four years and let Sheela and I, free for my PhD. I am equally thankful to my 4-year-old daughter who was living away from us with my mother, and for letting me do my work even after joining us recently

I would also like to thank laboratory technicians Andy Brook, Nick Skidmore, Ann Dawson, Balbir Bains, Keith Holding, Clare Murcott, Henrik Townsend and other technicians, who supported me throughout the project by providing me all the necessary equipments, materials and reagents. I would also like to extend my appreciation to the research administrator, Raman Kaur for being extremely helpful throughout the project.

Also, my sincere thanks go to my colleagues Vinu, Hoda, Prasanna, Peng, Patricia, GPeng, Sathish, Raji, Chandra and Lawrence for their support. I would also like to extend my gratitude to my friends outside of my University community for their support and contributions.

Last but not the least, I am thankful to the University of Wolverhampton for providing me with the scholarship to pursue my PhD, without this scholarship, my dream of achieving a PhD - an academic pinnacle of my life – would not have been possible.

June 2015

Wolverhampton, United Kingdom

Abstract

Tumour metastasis to the brain is a common and deadly development in certain cancers; 18-30% of breast tumours metastasise to the brain. The contribution that gene silencing through epigenetic mechanisms plays in metastasis to the brain is not well understood.

To identify epigenetic drivers of brain metastasis, a combined candidate gene screen using literature review, bioinformatics analysis of 450K methylation data of primary breast tumours from The Cancer Genome Atlas (TCGA) and Genome-wide methylation analysis of metastatic brain tumors that originated from primary breast tumours were carried out.

A candidate gene approach identified two genes (*BNC1* and *CCDC8*) dysregulated in breast to brain metastases (BBM) from a screen of 78 genes. Similarly, bioinformatic analyses of TCGA data identified *GALNT9* and an independent comparison of genome-wide methylation profiles in brain metastases identified 7 genes including non-coding RNA genes dysregulated in BBM. Taken together, these 10 genes identified are metastatic suppressor or promoter genes, which include novel regulatory elements non-coding RNA (ncRNAs) genes such as microRNAs, long intergenic non-coding RNAs (lincRNAs) or non-protein coding genes such as pseudogenes derived from their parental gene. Methylation analyses in BBM and their associated primary tumours from individual patients have revealed that identified genes are dysregulated either early or late in tumour evolution due to aberration in DNA methylation. In addition, methylation status of these genes in BBM correlates to serum DNA methylation in individual patients, which suggests that these genes could be used as a panel of prognostic markers or as therapeutic targets for BBM.

Contents

CHAPTER 1

Introduction

1.1 Cancer as a disease

1.1.1 Global incidence of cancer 1

1.1.2 Genetic and cellular basis of Cancer 3

1.2 The Role of Epigenetics in Cancer 4

1.2.1 The role of DNA methylation

1.2.1.1 Origin of DNA methylation and its maintenance in cells 4

1.2.1.2 The mechanisms of gene silencing and activation by DNA methylation 6

1.2.1.3 The role of DNA methylation in cancer 7

12.2 Role of histone modifications in cancer 10

1.3 Cancer Metastases

1.3.1 Dissemination of tumour cells from the primary tumours 15

1.3.2 Seed and soil theory of cancer metastases 18

1.3.3 Metastatic Cascades 19

1.3.4 Biological mechanisms underlying metastatic dormancy 21

1.3.5 Interaction between the microenvironment and the tumour cell 24

1.3.6 Role of metastatic suppressors in dormancy of disseminated tumour cells 25

1.4 Cellular and molecular basis of metastasis

1.4.1 Epigenetic regulation in cancer metastases 25

1.4.2 Role of angiogenesis in metastatic evolution 26

1.4.3 Role of tumour hypoxia in invasion and metastasis 28

1.4.4 Tumour microenvironment 29

1.4.5 Epithelial-Mesenchymal Transition (EMT) 30

1.5 The Brain metastases

1.5.1 Patterns of brain metastases 33

1.5.2 The brain microenvironment 34

1.5.3 Role of brain microenvironment in brain metastases 36

1.5.4 BBM 37

1.6 The biology of breast tumour

1.6.1 Primary breast tumour and its subtypes 40

1.6.2 Evolution of the metastatic breast tumour 43

1.6.3 Prognostic and predictive indicators for breast tumour recurrence

1.6.3.1 Tumor phenotype, patients' age/ethnicity and vascular/lymphatic status 43

1.6.3.2 Hormone receptor (ER/PR/HER2) Status 45

1.6.4 Mechanisms of dormancy of disseminated breast tumour cells 45

1.6.5 Detection of micrometastases and late occurrences 46

CHAPTER 2

Aims and Objectives

2.1 The Project Background 48

2.2 Identifying genes dysregulated in breast to brain metastases

2.2.1 A broad literature search to screen genes frequently methylated in one of the tumour types that metastasises to the brain 50

2.2.2 A broad literature search for screening of genes downregulated in EMT 52

2.2.3 Bioinformatic analysis of The Cancer Genome Atlas (TCGA) data 52

2.3 Investigation of methylation status of candidate genes and their expression in methylated in primary breast tumours and BBM 54

2.4 Study of metastatic potential of candidate genes silenced or downregulated in BBM due to promoter hypermethylation

2.4.1 Wound healing assay 57

2.4.2 Invasion assay 57

2.5 Identification of novel candidate genes by genome wide analyses of aberration in DNA methylation contributing to BBM 59

2.6 Investigate whether the genomic dysregulation in BBM is an early or a late event during the evolution of BBM in individual patients 59

2.7 Investigate whether the candidate genes identified (metastatic suppressor or metastatic promoter) genes could be used as prognostic biomarkers for BBM 59

CHAPTER 3

Materials and Methods

3.1 Patients and Samples

3.1.1 BBM samples 60

3.1.2 Matched primary breast tumours 60

3.1.3 Genomic DNA from primary tumours 60

3.1.4 Plasma/serum of patients with BBM 61

3.1.5 Breast cancer cell lines 62

3.2 Nucleic Acid Extractions and Preparation

3.2.1 Genomic DNA Extraction

3.2.1.1 Genomic DNA extraction from fresh frozen metastatic brain tumours 63

3.2.1.2 Genomic DNA extraction from Formalin fixed Paraffin Embedded (FFPE) tumour samples 64

3.2.2 Extraction of tumour free circulating DNA from patients' plasma 65

3.2.3 RNA extraction 65

3.2.4 Quantification of nucleic acid preparations 67

3.2.5 Protein Extraction and quantifications 67

3.3 DNA Methylation Analyses

3.3.1 DNA methyltransferase modification of DNA to create fully methylated positive controls 68

3.3.2 Bisulphite Conversion of DNA 68

3.3.2.1 Bisulphite conversion of genomic DNA from metastatic brain tumours 69

3.3.2.2 Bisulphite conversion of tumour free circulating DNA from patients' plasma 69
3.3.3 Combined Bisulphite and Restriction Analysis (CoBRA)
3.3.3.1 CoBRA Primer design 71
3.3.2.2 CoBRA PCR 72
3.3.4 Methylation Specific PCR (MSP) and Quantitative MSP (qMSP)
3.3.4.1 MSP Primer design 74
3.3.4.2 MSP in brain metastases and plasma 75
3.3.5 Quantitative Methylation Analysis by Bisulphite Sequencing
3.3.5.1 Purification of PCR Products 75
3.3.5.2 Ligation of PCR products into plasmid vector 76
3.3.5.3 Preparation of ampicillin containing (Amp+) agar plates and LB Broth 77
3.3.5.4 Transformation of PCR product ligated into the pGEM-T easy vector into the bacterial cells 77
3.3.5.5 Single colony PCR 78
3.3.5.6 PCR product clean-up for sequencing 78
3.3.5.7 Sequencing reaction 78
3.3.5.8 Ethanol precipitation following sequencing reaction 79
3.4 Expression Analysis of Selected Genes
3.4.1 Synthesis of complementary DNA (cDNA) from genomic RNA 79
3.4.2 Reverse transcription (RT) primer design 80
3.4.3 RT PCR 81
3.4.4 Quantification of gene expression 81
3.4.5 Western Blotting 81
3.5 Tissue Culture 84
3.5.1 Preparation of growth media 84
3.5.2 Revival of cell lines 84
3.5.3 Maintenance and passaging of cell lines 85
3.5.4 Cell counting using a Haemocytometer 85
3.5.5 Treatment of cell lines with 5-aza-2'deoxycitidine 86
3.5.6 Preparation of cell pellets 87

3.5.7 Freezing-down cell lines 87

3.6 Functional Analysis of genes/Proteins

3.6.1 RNA Interference (RNAi) assay and over-expression of genes in breast cancer cell lines 87

3.6.2 Migration assay 88

3.6.3 Invasion assay 89

3.7 Illumina Infinium Humanmethylation450k Data Analyses

3.7.1 Analyses of the 450K Methylation data from TCGA 90

3.7.2 Analyses of the 450K Methylation data for breast cancer brain metastases 92

3.8 Statistical Methods

3.8.1 Fisher's exact test 93

3.8.2 Kaplan-Meier analysis 93

CHAPTER 4

DNA methylation analyses to identify metastatic suppressor genes dysregulated in BBM

4.1 Introduction 95

4.2 Results

4.2.1 Screening of candidate metastatic suppressor genes in BBM 99

4.2.2 Identification of frequently methylated genes in metastatic brain tumours 100

4.2.3 Determination of methylation status of 20 genes in a cohort of unrelated primary breast tumours 109

4.2.4 Expression analysis of *BNC1* and *CCDC8* in metastatic brain tumours 119

4.2.5 Promoter methylation status of *BNC1* and *CCDC8* in brain metastases and associated primary breast tumours from individual patients 123

4.3 Discussion 127
4.4 Conclusions 139

CHAPTER 5

Identifying novel candidate genes dysregulated in BBM by analyses of the Genome wide 450K methylation array data from TCGA

5.1 Introduction 140
5.2 Results
5.2.1 Bioinformatic analyses of breast and lung tumours from TCGA 143
5.2.2 Methylation analyses of five BBM candidate genes 149
52.3 Methylation analyses of *GALNT9* in unrelated primary breast tumours 151
5.2.4 Expression analysis of *GALNT9* in metastatic brain tumours 157
5.2.5 Promoter methylation status of *GALNT9* in brain metastases and associated primary breast tumours from individual patients 160
5.2.6 Quantitative methylation analysis for *BNC1*, *CCDC8* and *GALNT9* 162
5.2.7 The Quantitative Expression analyses for *BNC1*, *CCDC8* and *GALNT9* 167
5.2.8 Methylation status of *BNC1*, *CCDC8* and *GALNT9* in metastatic and non-metastatic primary breast tumours from the TCGA 170
5.3 Discussion 179
5.4 Conclusions 184

CHAPTER 6

Functional analyses of *BNC1*, *CCDC8* and *GALNT9* to investigate their influence on the metastatic potential breast cancer cell lines.

1

6.1 Introduction 185

6.2 Results

6.2.1 Treatment of breast cancer cell lines with 5-Aza-2'deoxycytidine 187

6.2.2 Loss of *GALNT9*, *CCDC8* or *BNC1* expression increases the motility of breast cancer cell lines 189

6.2.3 Reduced expression of *BNC1*, *CCDC8* and *GALNT9* increases the invasive potential of breast cancer cell lines 194

6.2.4 Association of reduced expression of BNC1, *GALNT9* or *CCDC8* with poor relapse-free survival of patients 198

6.3 Discussion 201

6.4 Conclusion 207

CHAPTER 7

Genome wide 450K methylation analyses to identify genes dysregulated in BBM

7.1 Introduction 208

7.2 Results

7.2.1 Normalisation of TCGA data sets 209

7.2.2 Statistical analyses of the 450K-methylation data to identify genes differentially methylated in normal breast tissues and primary breast tumours, and BBM samples 210

7.2.3 Validation of methylation status of probes in BBM samples 225

7.2.4 Validation methylation status of the probes in Cohort of unrelated primary breast tumours 227

7.2.5 Methylation status of the genes in BBM samples and their corresponding primary breast tumours from individual patients 236
7.2.6 Investigating the methylation status of genes in patients' plasma to investigate if these genes could be used as prognostic markers for BBM 242
7.3 Discussion 247
7.4 Conclusion 257

CHAPTER 8

Final Discussion and Final Conclusion

8.1 Final Discussion 259
8.2 Final Conclusion 266

Appendix 268
References 304
Publications 329

Abbreviations

5-AZA-dC	5'-AZA-2'-deoxycytidine
ADH	Atypical Ductal Hyperplasia
BBB	Blood Brain Barrier
BBM	Breast to brain metastases
BM	Brain metastases
BME	Brain microenvironment
BN	Normal tissues adjacent to primary breast tumour
BP	Breast primary tumours
BRCA	Breast Invasive carcinoma
BTNW	Brain tumour North West
CGI	CpG island
CoBRA	Combined Bisulphite and Restriction Analysis
DMEM	Dulbecco's modified Eagle medium
DNMT	DNA Methyltransferase
DTC	Disseminated tumour cell
ECM	Extracellular Matrix
ECM	Extra Cellular Matrix
EGFR	Epithelial growth factor receptor
EMT	Epithelial-Mesenchymal Transition
ER	Oestrogen receptor
ERK	Extracellular signal regulating kinase
FBS	Foetal bovine serum

FEA	Flat epithelial atypia
FFPE	Formalin Fixed Paraffin Embedded
GABA	γ-acidic butyric acid
GEO	Gene expression Omnibus
HAT	Histone acetylase
HDAC	Histone deacetylase
HER2	Human epidermal growth factor receptor 2 (ERBB2)
HMT	Histone Methyl transferase
iPSCs	Induced pluoripotent stem cells
LBM	Lung to brain metastases
LUAD	Lung adenocarcinoma
MET	Mesenchymal to epithelial transition
MI	Methylation Index
MSP	Methylation Specific PCR
MSP	Methylation specific PCR
PR	Progesterone receptor
PCR	Polymerase chain reaction
RCC	Renal Cell Carcinoma
RT-PCR	Reverse Transcription PCR
SAM	S-adenosyl methionine
TCGA	The Cancer Genome Atlas
TSS	Transcription Start Site
USP	Unmethylation specific PCR
UTR	Unranslated region
VEGF	Vascular Endothelial Growth Factor

List of figures

Figure 1.1	The global cancer incidence and deaths	2
Figure 1.2	Mechanism of DNA methylation on gene silencing and genomic dysregulation of regulatory elements	9
Figure 1.3	The basic epigenetic mechanism such as DNA methylation and histone modifications in normal cells and in cancer.	12
Figure 1.4	A pattern of organotropism in cancer metastases	17
Figure 1.5	A linear and parallel progression model of cancer metastases	23
Figure 1.6	A process of EMT	32
Figure 2.1	The possible genomic alterations in cancer metastasis to the brain during metastatic cascade.	49
Figure 2.2	Screen of dysregulated genes in BBM	51
Figure 2.3	Rationale behind bioinformatics analysis of 450K methylation array data for primary breast and lung tumours from the TCGA	53
Figure 2.4	Graphical overview of methodologies used in screening and experimental validation of genes dysregulated in BBM	55
Figure 2.5	The experiments to investigate metastatic potential of breast cancer cells following the knockdown of candidate genes	58
Figure 4.1	Summary of screening methodology and experimental analyses to identify genes dysregulated in BBM.	99

Figure 4.2	Methylation frequency of candidate metastatic suppressor genes in BBM versus primary breast tumours	117
Figure 4.3	Methylation status of *BNC1* in BBM and unrelated primary breast tumours	119
Figure 4.4	Methylation status of *CCDC8* in BBM and unrelated primary breast tumours	120
Figure 4.5	Gel electrophoresis of RT-PCR of *BNC1* and *CCDC8* in BBM	123
Figure 4.6	Methylation status of *BNC1* and *CCDC8* in metastatic brain tumours and their corresponding originating primary breast tumours from individual patients	127
Figure 5.1	Bioinformatic analyses of TCGA data to screen genes unique for BBM and LBM	148
Figure 5.2	Methylation status of *GALNT9* in BBM.	152
Figure 5.3	Methylation status of *GALNT9* in unrelated primary breast tumours	157
Figure 5.4	Gel electrophoresis of RT-PCR of *GALNT9* in BBM	161
Figure 5.5	Methylation status of *GALNT9* in metastatic brain tumours and their corresponding originating primary breast tumours from individual patients	163
Figure 5.6	The promoter region/CpG islands of *BNC1*, *CCDC8* and *GALNT9*.	165
Figure 5.7	Bisulphite sequencing of individual alleles from tumours	167
Figure 5.8	Expression levels of each gene quantified in relation to the expression of β-actin	170
Figure 5.9	CoBRA primers designed for *BNC1*, *CCDC8* and *GALNT9* to	174

	investigate the methylation status of promoter region of each gene	
Figure 6.1	Global demethylation resulted in the reexpression of *BNC1, CCDC8* and *GALNT9* in breast cancer cell lines	192
Figure 6.2	RT-PCR to confirm Knockdown of *BNC1*, *CCDC8*, and *GALNT9* in breast cancer cell lines prior to functional assays	195
Figure 6.3	Wound healing assay showing increase in migratory potential of breast cancer cell lines due to loss of *CCDC8, BNC1* or *GALNT9* expression	196
Figure 6.4	Invasion assay showing increase in invasive potential of breast cancer cell lines due to reduced expression of *BNC1, CCDC8* or *GALNT9*	164
Figure 6.5	RT-PCR and Western blot to confirm knockdown of *BNC1*, *CCDC8* and *GALNT9* in breast cancer cell lines set up in parallel during migration and invasion assays	201
Figure 6.6	Kaplan-Meier analysis for *BNC1, CCDC8* and *GALNT9* via the prognoscan database	203
Figure 7.1	Methylation status of *HSPB9* in BBM and unrelated primary tumours	222
Figure 7.2	Analyses of the 450K-methylation array data to identify two classes of genes dysregulated in BBM	225
Figure 7.3	Representative examples of the probes, which are frequently methylated in BBM samples compared to normal breast tissues and primary breast tumours	228
Figure 7.4	Representative examples of the probes, which are infrequently	230

	methylated in BBM samples compared to normal breast tissues and primary breast tumours	
Figure 7.5	Methylation status of potential candidate metastatic suppressor genes in BBM versus unrelated primary tumours	237
Figure 7.6	Methylation status of potential candidate metastatic promoter genes in BBM versus unrelated primary tumours	239
Figure 7.7	Graphical representation of seven candidate genes differentially methylated in BBM compared to primary breast tumours and normal breast tissues	242
Figure 7.8	Methylation status (denoting early events) of genes in metastatic brain tumours and their corresponding originating primary breast tumours from individual patients	245
Figure 7.9	Methylation status (denoting late events) of genes in metastatic brain tumours and their corresponding originating primary breast tumours from individual patients	248
Figure 7.10	Methylation status of candidate prognostic markers (genes) in BBM versus tumour free serum DNA	253
Figure 8.1	Graphical representation of all identified candidate prognostic markers dysregulated in BBM and is differentially methylated in BBM compared to primary breast tumours.	270

List of tables

Table 3.1	The matched pairs of tumours: BBM, primary breast tumours and normal breast tissues from the individual patients	61
Table 3.2	A cohort of patients' serum and their corresponding BBM samples from individual patients	62
Table 3.3	Reagents used to set up nested and semi nested PCR reaction	73
Table 3.4	Touchdown PCR program	74
Table 3.5	Concentration of primary antibody and secondary antibody, exposure time and molecular weight of proteins BNC1, CCDC8 and GALNT9 analysed by western blotting	84
Table 3.6	A barcode identifiers of primary and lung tumours from TCGA to carry out bioinformatic analyses of 450K methylation array data	92
Table 4.1	Promotor methylation status of 42 candidate metastatic suppressor genes in BBM.	103
Table 4.2	Promotor methylation status of 36 candidate genes, which are down regulated in EMT.	105
Table 4.3	Genes frequently methylated in BBM samples and their functions	107
Table 4.4	Methylation status of candidate metastatic suppressor genes in primary breast tumours	112

Table 4.5	Promoter methylation status of candidate metastatic suppressor genes in a second cohort of BBM samples	115
Table 4.6	Clinical information of originating primary breast tumours and methylation status *BNC1* and *CCDC8*.	139
Table 5.1	Genes identified from TCGA data analyses that may contribute to BBM	149
Table 5.2	Methylation status of genes identified from TCGA data analyses in BBM	155
Table 5.3	Clinical information of metastatic primary breast tumours from TCGA	176
Table 5.4	Methylation status of *BNC1*, *CCDC8* and *GALNT9* in primary breast tumours from TCGA	177
Table 5.5	Methylation statuses of *BNC1*, *CCDC8* and *GALNT9* in non-metastatic breast tumours from TCGA	179
Table 5.6	Clinical information of originating primary breast tumours and methylation status *GALNT9*	185
Table 7.1	Genes identified from the 450K-methylation data analyses based on mean β value of all the samples.	219
Table 7.2	Methylation status of the genes in BBM and in a cohort of unrelated primary breast tumour samples	220
Table 7.3	The list of probes selected for laboratory analyses, which were hypermethylated in BBM samples compared to normal breast tissues and primary breast tumours	226
Table 7.4	The list of probes selected for laboratory analyses, which were hypomethylated in BBM samples compared to	227

normal breast tissues and primary breast tumours

Table 7.5 Methylation status of array analysis candidate genes in BBM as determined by CoBRA 233

Table 7.6 Methylation status the candidate genes in a cohort of unrelated primary tumours 235

Table 7.7 Methylation status of candidate prognostic markers in BBM and tumour free circulating DNA 252

CHAPTER 1

Introduction

1.1 Cancer as a disease

1.1.1 Global incidence of cancer

One in seven deaths in the world are caused by cancer (American Cancer Society, 2015); a disease comprised of more than 100 subtypes, these subtypes originate from most of the different cell types found in adults and children (Stratton *et al.*, 2009). The incidence of cancer is rising, factors such as ageing populations, urbanization and other lifestyle factors such as smoking, obesity and changes in reproductive patterns are believed to contribute to the rise in cancer cases globally (Torre *et al.*, 2015). In 2012, there were 6.7 million new cases of cancer in females and 7.4 million cases in males making a total of 14.1 new cases of cancer globally (Cancer Research UK, 2015). As of 2012, lung cancer remained the major cause of deaths in males globally; breast cancer is the second biggest cause of global deaths, the majority of these deaths occurring in women (Torre *et al.*, 2015). The cancer incidence and deaths by major cancer types in males and females are illustrated in figure 1.1.

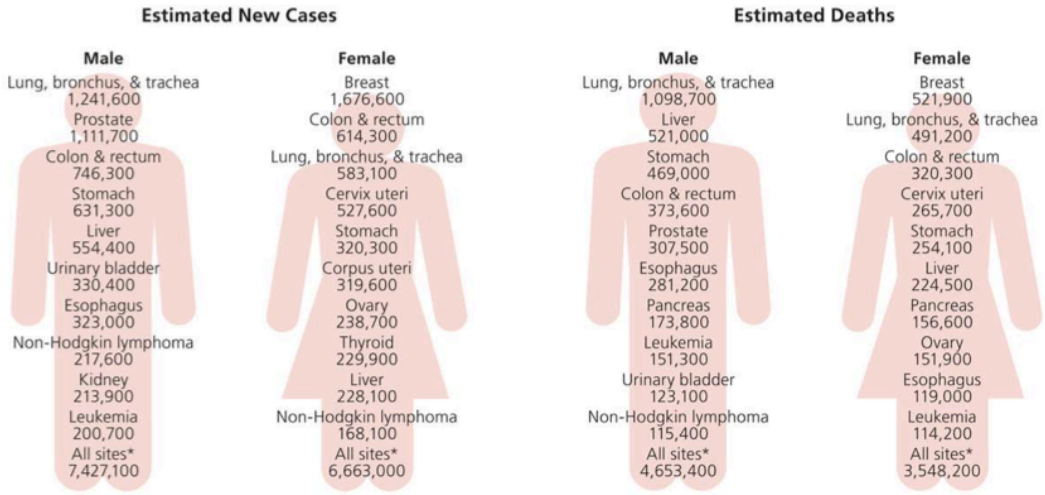


Figure 1.1: The global cancer incidence and deaths in 2012. The major cause of deaths by cancer worldwide is cancer of the lungs including bronchial and tracheal cancers in males and breast cancer in females. The types of cancer causing the most deaths in both men and women are lung, liver, breast, stomach, colon and rectal, oesophagus, pancreas and leukaemia. The figure is taken from (Torre *et al.*, 2015).

1.1.2 Genetic and cellular basis of Cancer

Cancer is characterised by uncontrolled cellular proliferation (American Cancer Society, 2015) followed by metastasis to different organs. The role of gene aberrations in cancer was first suggested when unusual chromosomal abnormalities were observed in dividing cancer cells under the microscope in the late nineteenth century (Hansemann, 1890). This observation paved the way to the conclusion that a tumour is a mass of abnormal clones with aberrations in genetic material (Hungerford & Nowell, 1962). More recently, different types of genetic abnormalities such as mutations, translocations, and genomic rearrangements became evident and specific aberrations such as mutations in specific positions such as G to T transition (as in HRAS gene) was identified (Reddy *et al.*, 1982). Hence, cancer arises due to a series of genetic aberrations that follows the process of Darwinian evolution within the microenvironment among cellular populations. These genetic alterations provide selective advantages to the tumour cells enabling them to grow and proliferate more effectively compared to their neighbouring cells (Stratton *et al.*, 2009). The genetic aberrations that occur in cancer cells could either be cancer drivers or passengers; a driver mutation contributes to the transformation of normal cell into a cancer cell whereas a passenger mutation does not obviously contribute to tumorigenicity (Schinzel & Hahn, 2008). In addition to classical mutation, epigenetic programing is also dysregulated in tumours, such that errors result in aberrations in chromatin structure and gene expression patterns (See section 1.2). Therefore, the origin of cancer or tumorigenesis is a result of a series of genetic and epigenetic changes that lead to defects of cellular growth, proliferation and differentiation (Schinzel & Hahn, 2008).

1.2 The Role of Epigenetics in Cancer

The term epigenetics, which refers to the heritable changes in the pattern of gene expression not directly mediated by alterations in the primary nucleotide sequence, was first used by Conrad Waddington to explain the interactions between the genes and their environment (Jones & Baylin, 2007; Dwivedi *et al.*, 2011). The study of epigenetics has developed into one of the most promising fields of biomedical research from its beginning in 1940, epigenetic regulation is now known to be associated with many human diseases including human cancers (Rodriguez-Paredes & Esteller, 2011). Various crucial cellular mechanism that regulate gene silencing such as imprinting and X-chromosome inactivation are essential during development and such mechanisms are required for stable cell proliferation and growth (Jaenisch & Bird, 2003). Epigenetic mechanisms such as DNA methylation, RNA-mediated gene silencing and histone modification lead to changes in chromatin dynamics that mediate cellular reprogramming necessary for development and cellular identity (Herceg & Vaissiere, 2011). Dysregulation of these epigenetic processes contribute to aberrations in gene expression giving tumour cells a selective growth advantage that contributes to cancer development (Tsai & Baylin, 2011).

1.2.1 The role of DNA methylation

1.2.1.1 Origin of DNA methylation and its maintenance in cells

DNA methylation refers to the addition of methyl groups (CH_3) to cytosine residues resulting in the formation of 5-methyl cytosine (5mC) (Bird, 2002). DNA methylation is primarily associated with silencing of the genes, which are necessary to be silenced for normal cellular development. Such gene silencing is important during embryonic

development (Okano *et al.*, 1999), genomic imprinting (Li *et al.*, 1993), inactivation of X-liked genes (Loyn, 1961) as well as silencing of DNA sequences such as mobile genetic elements (retroposons or transposons) and endogenous retroviruses (Kass *et al.*, 1997). DNA methylation in mammals (cytosine methylation) is carried out by DNA methyltransferases (DNMT) 1, 2 and 3 (Bird, 2002; Schaefer *et al.*, 2008) among which DNMT1 is the most abundantly present DNA methyltransferase that maintains methylation during DNA replication by methylating hemimethylated CpG dinucleotides (Okano *et al.*, 1999). Similarly, tRNA methyltransferase activity is maintained by DNMT2 (Schaefer *et al.*, 2008) whereas DNMT3 family carries out *de novo* methylation during embryonic development. The DNMT3 family includes DNMT3A and DNMT3B; DNMT3A mediates genomic imprinting through methylation during gametogenesis whereas both DMNT3A and DNMT3B mediate methylation on repetitive sequences (Okano *et al.*, 1999). CpG dinucleotides are present throughout the mammalian genome; however, the presence of CpG rich regions at the 5'end of the gene, which are termed CpG islands (CGI), is one of the remarkable features of the eukaryotic genome. By definition, CGI is a region with more than 50% CGs, covering a minimum length of 200 base pairs (Gardiner-Garden & Frommer, 1987). It has been observed that the majority of the CGIs around promoter region of the genes are unmethylated during embryonic development and a significant portion becomes methylated during development resulting into stable gene silencing (Antequera & Bird, 1993). In addition, the majority of CGIs outside the promoter region and other CpG dinucleotides remain methylated in normal somatic cells (Deaton *et al.*, 2011). In germ cells and peri-implantation stages of development in mammals, the DNA methylation is erased and reprogrammed by DNMT3A and DNMT3B (*de novo* methyltransferases), which are also needed to maintain original patterns of DNA methylation during early

development (Monk *et al.*, 1987; Kafri *et al.*, 1992). Similarly, the maintenance methylation is mediated between the successive generations by DNMT1 that effectively copies the methylated and unmethylated CpGs in the newly formed daughter DNA strands during semiconservative replication of DNA (Pradhan *et al.*, 1999). Therefore, in normal somatic cells, *de novo* methylation established by *de novo* methyltransferases DNMT3A and DNMT3B is maintained by maintenance methyltransferase DNMT1 between the successive generations.

1.2.1.2 The mechanisms of gene silencing and activation by DNA methylation

There are two principal mechanisms by which DNA methylation mediates transcriptional gene silencing. First, the repression of transcription is brought about by preventing binding of transcriptional factors on gene promoters as exemplified by transcription factor N-Myc binding on CGI of its target gene (Perini *et al.*, 2005). Second, gene silencing due to methylation is mediated by recruiting methyl DNA binding proteins (MDBPs). The Kaiso like family of proteins, the SRA (SET- and RING-associated) domain proteins and methyl binding domain family of proteins (MDBs) are the three classes of MDBPs, of which MBDs are the most common and biggest players in gene repression (Yoon *et al.*, 2003). The recruitment of MDBPs results in recruitment of additional repressor proteins that in turn recruit histone modification complexes such as histone deacetylases (HDACs) leading to chromatin modifications (Zhang *et al.*, 1999). Four classes of MDBs namely MDB1, MDB2, MDB3 and MDB4 (Roloff *et al.*, 2003) and another two classes of MBDs MeCP1 and

MeCP2 (Lewis *et al.*, 1992) have been identified that play roles in transcriptional repression in a similar manner (Cihák, 1974).

1.2.1.3 The role of DNA methylation in cancer

Aberrant DNA methylation change is one of the most widely studied epigenetic mechanisms in cancer, characterised either by global DNA hypomethylation or hypermethylation or by hypomethylation at a localised region of the gene.

A global loss of methylation or hypomethylation can lead to the activation of repeat sequences, which otherwise remain silenced in the genome, resulting in chromosomal rearrangement and genomic instability (Xu *et al.*, 1999). It has been observed that the global hypomethylation in repeat elements such as Alu and LINES elements has led to genomic rearrangement and activation of these transposable genetic elements (Weisenberger *et al.*, 2005). In addition, analysis of chromosome 11 in mice that contain a hypomorphic *Dnmt1* allele, an allele with a partial loss of Dnmt1 function (Dnmt+/-) and single copy of *TP53* and *NF1* has shown genomic instability due to hypomethylation (Eden *et al.*, 2003).

In addition, gene specific hypomethylation can occur, activating proto-oncogenes which otherwise would remain silenced (figure 1.2). A study on salivary gland adenoid cystic carcinoma has showed that 8 genes were overexpressed due to loss of methylation (Shao *et al.*, 2011). Promoter hypermethylation of common tumour suppressor genes (figure 1.2) and their subsequent inactivation or down regulation was observed in many cancers such as *RB1* in retinoblastoma (Greger *et al.*, 1989; Sakai *et al.*, 1991), *VHL* in renal cell carcinoma (Herman *et al.*, 1994) and *CDKN2A* (*cyclin-dependent kinase*

inhibitor 2A) in multiple tumour types (Herman *et al.*, 1995). In several cases, downregulation or complete silencing of the tumour suppressor gene is the principal mechanism of dysregulation in cancer, as is evident in *RASSF1A* inactivation in various cancers such as lung, breast, glioma, colorectal and Renal Cell Carcinoma (RCC) (Hesson *et al.*, 204; Dammann *et al.*, 2000; Dammann *et al.*, 2001; Morrissey *et al.*, 2001; van Engeland *et al.*, 2002) and *HIC1* inactivation in leukemia and breast cancer (Issa *et al.*, 1997; Fujii *et al.*, 1998; Melki *et al.*, 1999)**.** Similarly, hypermethylation and downregulation of tumour suppressor gene may be a second hit as evident in *CHD1* hypermethylation in various cancers (Grady *et al.*, 2000). Aberrant DNA methylation has now been identified in hundreds of genes and in most cancer types such as childhood acute lymphoma (Dunwell *et al.*, 2009b), glioma (Kim *et al.*, 2006; Patel *et al.*, 2008), and other CNS tumours (Muhlisch *et al.*, 2006; Margetts *et al.*, 2008), bone (Rao-Bindal & Kleinerman, 2011), breast (Miyamoto *et al.*, 2003; Yuan *et al.*, 2003; Thakur *et al.*, 2007; Moelans *et al.*, 2011; Park *et al.*, 2011; Zeng *et al.*, 2012), lungs (Miyamoto *et al.*, 2003; Shames *et al.*, 2006a; Hsu *et al.*, 2007; Han *et al.*, 2009; Paliwal *et al.*, 2010), colorectal (Miyamoto *et al.*, 2003), pancreas (Miyamoto *et al.*, 2003), renal (Morris *et al.*, 2005; Morris *et al.*, 2008; McRonald *et al.*, 2009; Morris *et al.*, 2010; Morris *et al.*, 2011) and melanoma (Hoon *et al.*, 2004; Marini *et al.*, 2006; Liu *et al.*, 2008; Koga *et al.*, 2009; Schinke *et al.*, 2010; Bonazzi *et al.*, 2011). The large number of genes dysregulated in cancer may contribute to tumorigenicity (driver events), however some are likely to be passenger events dysregulated as a consequence of genome-wide epigenetic dysregulation.

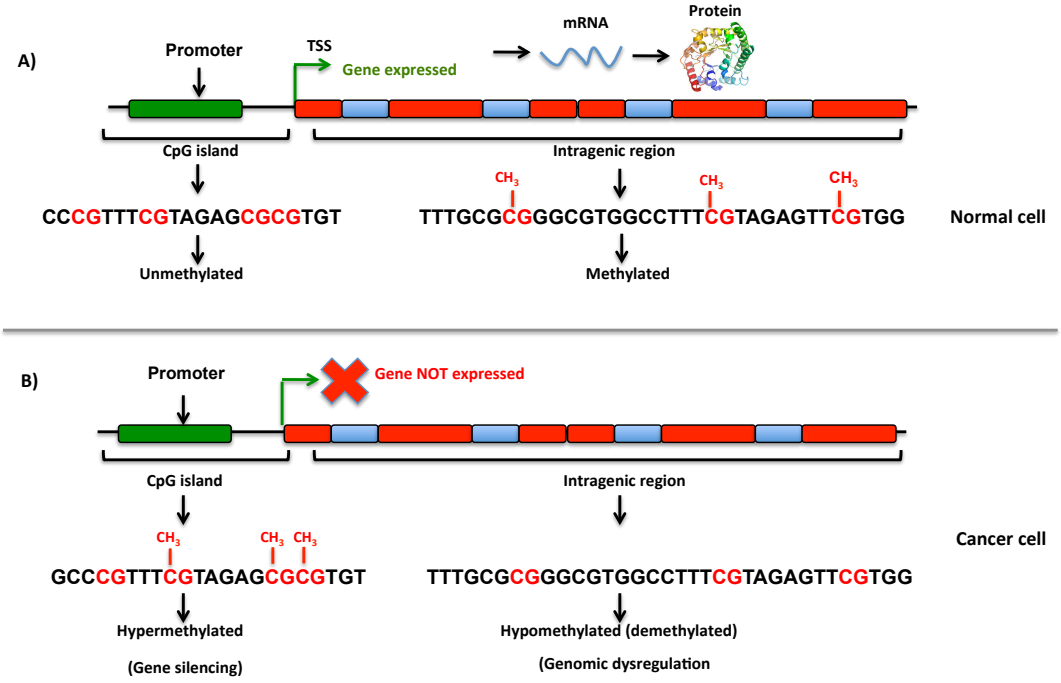


Figure 1.2: Mechanism of DNA methylation on gene silencing and genomic dysregulation of regulatory elements. CpG island promoter regions are generally unmethylated in tumour suppressor genes in normal cells (A), which when hypermethylated (B), results in the transcriptional silencing in cancer. In contrast, tumour promoter genes are kept silenced due to promoter methylation in normal cells (B), which when hypomethylated, results in the transcriptional activation in cancer cells (A). Similarly, aberration of DNA methylation (hypomethylation/demethylation) of intragenic region of regulatory elements results in genomic instability.

1.2.2 Role of histone modifications in cancer

A eukaryotic genome consists of double stranded DNA complexed with histones forming a repeating subunit called the nucleosome, which further condenses to form chromatin. Each nucleosome consists of two subunits of each histone H2A, H2B, H3 and H4 forming a complex with another linker histone H1 (Kornberg, 1974). In a normal cell, the chromatin being a highly packaged, condensed architecture, either remains in a repressive state (heterochromatin) or in a less condensed state (euchromatin), which is accessible to the transcriptional machinery for gene expression (Grunstein *et al.*, 1995).

Histone modification refers to the various types of histone modifications that are reported to be involved in normal cellular physiology and in cancer. These include acetylation, methylation, phosphorylation, ubiquitination, sumoylation, ADP-ribosylation, deamination, proline isomerism and propionylation (Sawan & Herceg, 2010) which either keep chromatin in a repressive or active state regarding transcriptional regulation. A great deal of work is ongoing to elucidate the complex nature of these modifications (Thurman *et al.*, 2012). However, the role of histone acetylation and methylation in cancer is reasonably well understood: In a normal cell physiology, one or more types of histone modifications maintain the chromatin state either as accessible (euchromatin) or less accessible state (heterochromatin) to transcriptional machinery to promote or to inhibit expression of genes (Quina *et al.*, 2006). For instance, chromatin can be made accessible to DNA by histone acetylation removing the positive charge at lysine residues on the histone tail that weakens the DNA-histone or nucleosome- nucleosome interactions; HATs and Histone deacetylases HDACs are the two enzymes governing histone acetylation and deacetylation

respectively, which are reported to be involved in cancer (Leroy *et al.*, 2013) (figure 1.3).

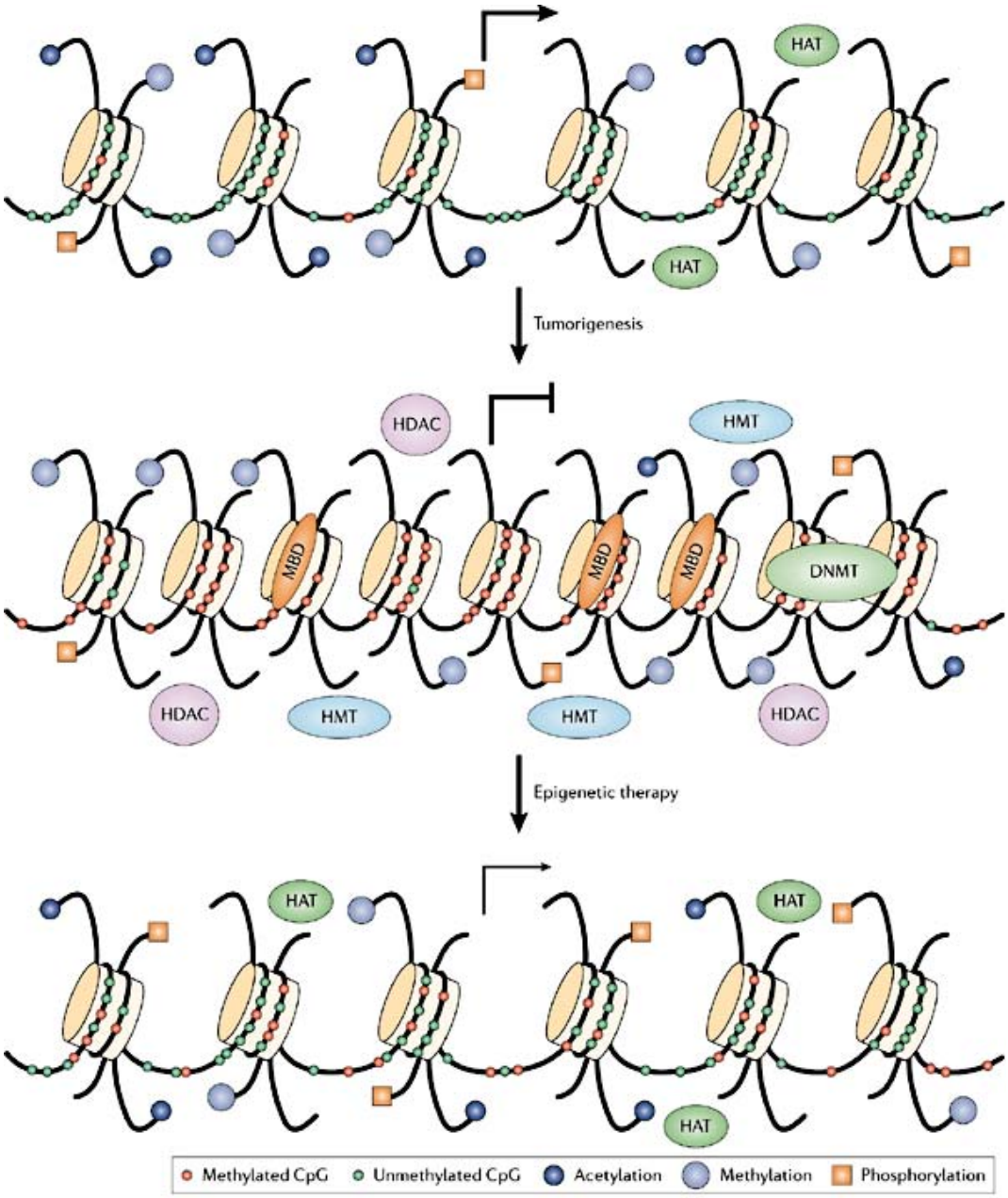


Figure 1.3: The basic epigenetic mechanism such as DNA methylation and histone modifications in normal cells and in cancer. In normal cell, the majority of CpGs in DNA are unmethylated and the chromatin is transcriptionally active (euchromatin) due to histone acetylation governed by histone acetylates (HATs) where as in cancer cells the majority of CpGs are methylated due to DNA

methyltransferases (DNMTs), that recruits methyl binding proteins (MBDs). MBDs further recruit histone deacetylases (HDACs) turning chromatin into transcriptionally inactive (heterochromatin). In addition, there are other modifications regulating chromatin dynamics such as histone methylation governed by histone methyltransferases (HMTs) and phosphorylation. The epigenetic modifications leading to gene silencing due to DNA methylation and histone modification can be reversed by epigenetic drugs (therapy) such as demethylating agents that could reactivate the silenced genes (The figure has been taken from Yoo and Jones, 2006) (Yoo & Jones, 2006).

Dysregulation of various HDACs has been reported in various cancers as exemplified by the upregulation of HDAC2, 3, 6, 8 and 7 in childhood Acute Lymphoblastic Leukemia (ALL) (Moreno *et al.*, 2010), overexpression of HDAC1, 2 and 3 in ovarian cancer (Hayashi *et al.*, 2010), overexpression of HDAC6 in breast cancer and oral squamous cell carcinoma (Sakuma *et al.*, 2006) and over expression of HDAC2 in numerous cancers (Jin *et al.*, 2008; Adams *et al.*, 2010; Langer *et al.*, 2010; Mutze *et al.*, 2010) .

Another crucial type of histone modification, histone methylation is catalyzed by histone methyltransferases that occurs at an arginine which is tri-methylated or at lysine residues which is mono-, di- or tri- methylated. These histone modifications marks are associated with either repressive or active chromatin states. For instance, mono methylation on lysine resides on different histone subunits are associated with transcriptional activation (H2BK5, H3K4, H3K9, H3K27, H3K79 and H4K20) (Barski *et al.*, 2007; Benevolenskaya, 2007; Steger *et al.*, 2008). Similarly, di- or tri-methylations on lysine residues are associated with either activated (H3K4m2, H3K4m3 and H3K79m2) (Koch *et al.*, 2007; Steger *et al.*, 2008) or repressive (H2BK5m3, H3K9m2, H3K9m3, H3K27m2, H3K27me3) (Barski *et al.*, 2007; Rosenfeld *et al.*, 2009) state. Interestingly, tri-methylation at lysine residues on H3 (H3K79m3) is associated with either as active (Steger *et al.*, 2008) or repressive state (Benevolenskaya, 2007), which could be due to the activity of different effector proteins. Methylation at arginine residues on H3 and H4 (H3R17, H3R23 and H4R3) is associated with transcriptional activation (Berger, 2007). An example of dysregulation of a histone demethylase is *LSD1* (*lysine (K)-specific demethylase 1A*) upregulated in bladder cancer (Hayami *et al.*, 2011), estrogen receptor negative breast cancer (Lim *et al.*, 2010) and neuroblastoma (Schulte *et al.*, 2009). In addition, evidence for the importance of histone

modifications is provided by mutations in *PBRM* in renal cell carcinoma (Varela *et al.*, 2011a) and *ARID1A* in ovarian cancer (Jones *et al.*, 2010). It has been observed that inhibition of *LSD1* results in growth inhibition of cells where as its upregulation is associated with upregulation of various growth regulatory genes as well as the genes involved in chromatin remodeling (Lim *et al.*, 2010; Hayami *et al.*, 2011).

1.3 Cancer metastases

1.3.1 Dissemination of tumour cells from the primary tumours

Cancer metastases refers to the formation of new tumours in a distant site due to the dissociation of tumour cells from the primary tumour (Chambers *et al.*, 2002). The majority of the cancer related deaths (90%) is credited to metastases to distant organs (Nguyen & Massague, 2007; Rodenhiser, 2009) which, to a greater extent, is due to the resistance to the existing therapy that can eradicate most primary tumours (Valastyan & Weinberg, 2011; Neman *et al.*, 2014). The ability to detect metastatic tumours early, as opposed to once they have metastasised, appears to be a difficult hurdle to overcome in the treatment of disease (Chambers *et al.*, 2002). The complex nature of the interactions between Disseminating Tumour Cells (DTCs) and microenvironment could drive treatment resistance (Valastyan & Weinberg, 2011). It is well established that metastatic cells disseminate from the primary tumours by invading the lymphatic vessels and then only some cells with aggressive metastatic potential are capable of entering in to the blood circulation (intravasation) (Fernández-Periáñez *et al.*, 2013). In this case, metastatic colonisation may take in different organs from the same primary tumour (synchronous seeding) or spread metastatic tumour cells from one secondary site to another may occur (metachronous seeding) (Fernández-Periáñez *et al.*, 2013). Occasionally the metastatic tumour cells return back to the primary site, this is termed

self-seeding (Nguyen & Massague, 2007). The presence of undetectable micrometastases or latent metastatic lesions running parallel to the malignant progression of primary tumours or their occurrence a long time after primary tumours diagnosis makes prognosis and treatment of metastatic diseases difficult (Holmgren *et al.*, 1995). In addition, it has been observed in many instances that the development of the malignant tumours from the adenoma (adenocarcinoma) is a slower process than the development of the metastatic tumours from it, which suggests that the metastatic tumour formation takes place in parallel with the malignant tumour (Klein, 2009).

Metastases exert an organ specific pattern in terms of final colonization in distant organs (Figure 1.4) such as breast tumours often metastasise to bone, lungs, brain, liver and more infrequently to adrenal glands (Chambers *et al.*, 2002).

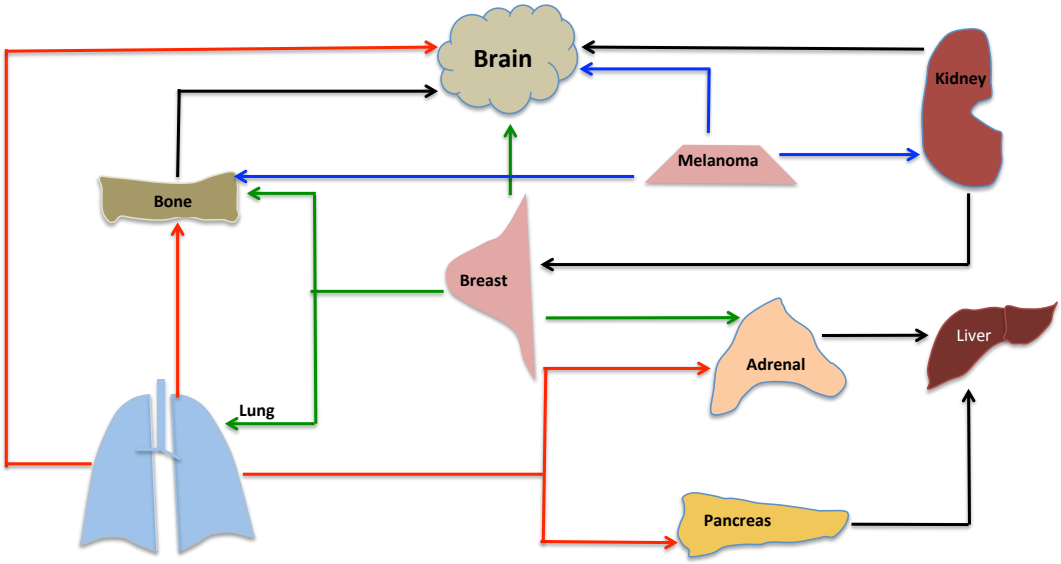


Figure 1.4 Metastases exhibit a pattern of organotropism, for instance, breast cancer metastasises frequently to the bone, lung, brain and liver and less frequently to adrenal glands (green arrows) whereas lung tumours metastasise to bone, brain, adrenal glands and pancreas (red arrows). Similarly, melanoma metastasises to brain, bone and kidney (blue arrows) whereas both adrenal glands and pancreas metastasises to the liver (black arrows).

Cancer metastasis always follows a multistep process regardless of the distant organs involved for metastatic evolution (Rodenhiser, 2009; Talmadge & Fidler, 2010; Cock-Rada & Weitzman, 2013) in which a tumour cell invades the extracellular matrix, disseminates to the blood stream (intravasation), survives in the circulation and invades (extravasation) the distant tissues and finally evolves (colonisation) as a metastatic lesions (Chambers *et al.*, 2002; Minn *et al.*, 2005b; Hu *et al.*, 2009; Rodenhiser, 2009; Zhang & Yu, 2011). Colonisation of DTCs in a distant sites consists of the metastatic events such as extravasation in the new niche, formation of micrometastases, survival in a new environment, induction of a latent period and evolution into a macrometastases (Vanharanta & Massague, 2013). Cancer can progress into distance sites of the body and may undergo a latent period for many years before developing into a more deadly and recurrent metastatic disease (Aguirre-Ghiso, 2007).

1.3.2 Seed and soil theory of cancer metastases

A British surgeon Stephen Paget in 1889 (Paget, 1889), after examining 700 cases of patients with breast cancer, postulated that cancer metastasis is not a random incident. Rather, it is the results of possible interaction between the tumour cells (seed) with the distant organs (soil) to which the tumour cell metastasises (Chambers *et al.*, 2002; Talmadge & Fidler, 2010; Zhang & Yu, 2011). The seed and soil hypothesis further clarifies that the seed may only be able to grow in the right soil even though it may spread in varying directions (Fidler & Poste, 2008; Talmadge & Fidler, 2010), in which a seed may be a progenitor cell, a cancer stem cell (CSC), metastatic cell or a tumour initiating cell that may be able to grow only in the suitable distant microenvironment (soil) (Talmadge & Fidler, 2010). Therefore, in addition to the neoplastic cells,

metastasis depends on the nature of the microenvironment to which metastatic cells tend to colonize (Fidler, 2003). In this regard, metastasis is a result of interaction between the tumour cells and their microenvironment including various factors such as angiogenesis, migration, invasion and cell proliferation (Fidler, 2003). This implies that the seed and soil theory, in principle, consists of the microenvironment of the distant organ and a heterogeneous population of neoplasm (primary tumour). In these regards, only some of the cells from each subpopulations may be able to complete some of the initial stages of the metastatic cascades (Klein, 2009). Therefore, it is possible that only very few cells, derived from a single cell, may be able to eventually colonise the distant niche which suggest that the metastatic tumour is of clonal origin and derived from a single tumour cell (Talmadge & Fidler, 2010). The cross talk between the soil (the microenvironment) and the homeostatic mechanism (Mendoza & Khanna, 2009) results into the metastatic development that when a seed (tumour cell) metastasises into a new and unique biological microenvironment (soil) *i.e.* specific distant organs (Paget, 1889; Talmadge & Fidler, 2010).

1.3.3 Metastatic cascades

In order to colonize distant sites, tumour cells from the primary tumours need to acquire genetic and epigenetic changes to overcome the sequential events, which are termed as the metastatic cascades. The first step in the metastatic cascades is a local invasion in which a cancer cells should be able to invade extracellular matrix (ECM), and penetrate (intravasate) the blood vessel (Chambers *et al.*, 2002; Coomber *et al.*, 2003) The blood microenvironment is different to the original primary tumour environment; cancer cells need to acquire the ability to resist the selective pressure of the blood vasculature such

as lack of adhesion, unfavourable conditions and immune responses. In the blood vasculature, it is necessary for tumour cells to adhere to the endothelial surface in the lumen of the blood vessels from which they migrate (extravasate) into the target organs (Coomber *et al.*, 2003). Tumour cells may remain dormant for several years attached to the surface of the blood vessel or the lymph vessel until they acquire additional layers of genetic and epigenetic dysregulation necessary to resume their progress, leading to the eventual colonisation of a new niche and the continued evolution into a distant site micrometastases (Nguyen *et al.*, 2009). Due to the sequential selection pressure, each stage in the metastatic cascade is rate limiting and the process of metastasis is regarded as inefficient through which only a few tumour cells will succeed to complete the process (Talmadge & Fidler, 2010). It has been demonstrated that some of the clones in a genetically and phenotypically heterogeneous population of primary tumour have a metastatic phenotype or factors influencing metastatic potential such as migration, invasion, survival, proteolysis and angiogenesis (Joyce & Pollard, 2009). Cancer metastasis, therefore, is a non-random, highly specialised process as postulated in the seed and soil hypothesis, in which the tumour cells capacity to proliferate in a specific distant site is either predetermined before the initial step or is determined during the process of metastatic cascade (Coomber *et al.*, 2003 ; Psaila & Lyden, 2009). In this regard, the metastatic phenotype, which is a result of the genetic and epigenetic dysregulation necessary for metastatic competence is similar to a higher-grade malignant phenotype due to the stepwise progression through a metastatic cascade (Nguyen & Massague, 2007). Therefore, the genetic or epigenetic changes responsible for the metastasis could possibly be predictable either in primary tumours as an early event or be able to be detected as the late event after the tumour progresses through the circulating tumour cells in the blood vasculature (Psaila & Lyden, 2009). Some of the

predictable early markers for tumour metastases include p53 mutations in advanced colorectal cancer (CRC) and amplification of *HER2* in metastatic breast cancer (Coomber *et al.*, 2003). However, it is likely that many more are yet to be identified.

1.3.4 Biological mechanisms underlying metastatic dormancy

A cancer cell, which has escaped from the primary tumour is regarded as a disseminated tumour cell, which after leaving the primary tumour site may undergo dormancy for years, some remain dormant for decades (Klein, 2009). The treatment of the primary tumour may lead to the DTCs persisting in the body as minimal residual tumours (Aguirre-Ghiso, 2007), which, with a favourable microenvironment, may develop into metastases. Another possibility is a dissemination of primary tumour cells to the distance site before the primary tumour reaches a state of malignancy in which a metastatic tumour grows in parallel with the primary tumour (Rocken, 2010). The genetic dysregulation in primary tumours is thought to provide genetic fitness or selective advantages to the tumour cells to metastasise as an early or a late event of the cancer progression (Aguirre-Ghiso, 2007; Cock-Rada & Weitzman, 2013). In an early event (parallel progression model), the successive rounds of genetic dysregulation provides the DTCs with the ability to exhibit more aggressive phenotypes to grow as metastatic tumours along with the primary tumours (Klein, 2009) (Figure 1.5A) whereas in late event (linear progression model), the tumours cells acquire genetic changes or epigenetic changes at a later stage of the primary malignant tumours to be able to invade the surrounding cells to enter into the metastatic cascades (Aguirre-Ghiso, 2007; Klein, 2009) (Figure 1.5B). The long time periods for a tumour cells to proliferate and to grow in to a metastatic lesions after the disseminations explains the period of

dormancy (Rocken, 2010), however, the mechanism of dormancy in different cancers is still elusive. Hence, in all cancers, the period of dormancy is attributed to the slow and steady accumulation of genetic aberrations such as loss of *TP53*, *RB1*, *P16*, *RAS*/*BRAF* activating mutations, *ERBB2* amplifications *etc*. required for transformation and immortalization of the tumour cells during carcinogenesis (Aguirre-Ghiso, 2007).

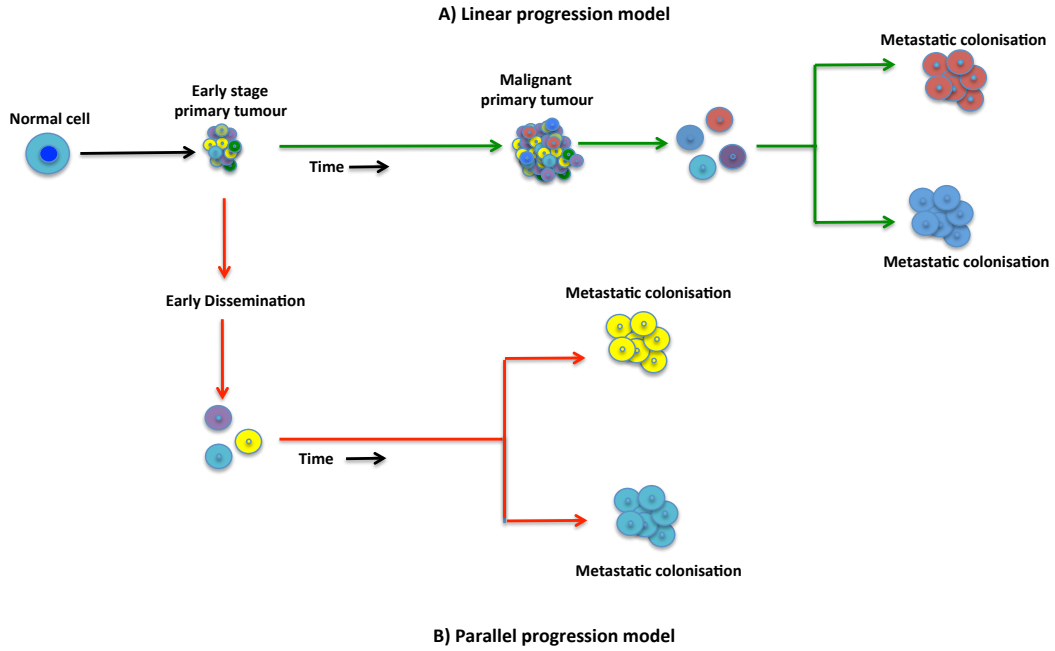


Figure 1.5: The two models of metastatic colonization. In a linear progression model (A), primary tumours progresses into a more aggressive malignant form from which the tumour cells disseminated to colonise distant organs whereas in a parallel progression model (B), tumour cells may disseminate early from early stage primary tumours and grow into a micrometastases along with the progression of the malignant primary tumour eventually forming a metastases.

1.3.5 Interaction between the microenvironment and the tumour cell

The new niche encountered by an invading tumour cell is not a natural environment to grow and proliferate. Tumour cells may be able to grow if the niche or a microenvironment is conductive for their growth. However, an unaccommodating microenvironment may significantly hinder the tumours growth capacity (Aguirre-Ghiso *et al*., 2004). The choice between growth arrest and proliferation of the tumours is decided by the deregulation of the interaction between the tumours and their microenvironment. Tumours cells may remain dormant if they have a capacity to remain in a differentiated state (Aguirre-Ghiso, 2007). Cross talk between the tumour cell and the DTCs was observed in head and neck carcinoma where the growth of the tumour was stimulated by metastasis associated urokinase receptor (uPAR) by activating α5β1 integrins (Gohji *et al*., 1997). This complex propagates mitogenic signals by an association with RAS and extracellular signal regulating kinase (ERK) by recruiting focal adhesion kinase (FAK) and EGF receptor (EFGR) (Aguirre-Ghiso, 2007). Similarly, inhibition of *RAF-MEK-ERK* signaling pathway by ERK signaling *in vivo* resulted in dormancy of tumour cell and G0-G1 arrest of the dormant cells (Aguirre-Ghiso *et al*., 2004). Likewise, the growth arrest or the dormancy of the tumour cells is favored by activation of *P38 MAPK* signaling due to the disruption of uPAR. Hence, high *ERK* activity contributes to sustain tumour growth where as the *P38* activation contributes to tumour dormancy by upregulation of *P53* and downregulation of c-*JUN* (Adam *et al*., 2009).

1.3.6 The role of metastatic suppressors in dormancy of disseminated tumour cells

Over 99% of DTCs are thought to die in a new microenvironment; the main hurdle that DTCs must overcome to enable their evolution into metastatic tumours is the deactivation of tumour suppressor genes (Aguirre-Ghiso, 2007). Tumour suppressor genes inhibit metastases by promoting apoptosis. For example, *KSS1* is thought to inhibit metastases by elongating the dormancy of DTCs, *CD82*, a tetraspanin protein inhibit metastases of mouse melanoma cell by binding to *CD234* (Aguirre-Ghiso, 2007; Rodenhiser, 2009). In addition to melanoma metastases, a reduced level of *KSS1* is implicated in breast cancer brain metastases (Rodenhiser, 2009). Similarly, RAS signalling is inhibited by *NM23*, which is expressed in non-metastatic cells. *MKK4* is silenced in both prostate and ovarian cancer by activation of *JNK* and *P38* respectively (Chambers *et al.*, 2002). Interestingly, a study using a mouse model has shown that the metastatic efficacy of the host depends on F1 parents suggesting the role for the genetic make up of the host on the ability to supress metastases (Crawford *et al.*, 2006; Aguirre-Ghiso, 2007).

1.4. Cellular and molecular basis of metastasis

1.4.1 Epigenetic regulation in cancer metastases

The success of the metastasising cell to complete the metastatic cascades leading to eventual colonisation depends on the interactions of tumour cells with host microenvironment at various stages of metastatic evolution. In these regards, the epigenetic regulation of microenvironment as well as the epigenetic dysregulation in

tumours cells contributes significantly to metastatic evolution (Coomber *et al.*, 2003). The basic epigenetic mechanism includes DNA methylation and histone modifications in which the various post-translational modification such as methylation and acetylation regulate chromatin dynamics, activating and inhibiting of gene expression (Rodenhiser, 2009). The process involves recruitment of enzymes for histone modification, which, either promotes or removes such modifications (Rodenhiser, 2009; Cock-Rada & Weitzman, 2013). A large number of histone modifying enzymes have been identified which regulate the metastatic phenotypes in primary tumours through their target genes (Cock-Rada & Weitzman, 2013). The epigenetic regulatory networks includes changes in the gene expression due to alterations in methylation, acetylation, and other processes which further contributes to various processes such as ECM degradation, post translational activation, changes in the adhesive properties of the cells, activation of coagulation signalling and activation of other enzymatic environment contributing to metastases (Coomber *et al.*, 2003). In addition, various factors such as growth factor availability, hypoxia, hormonal influences and metabolic stress, composition of ECM, inflammatory reaction and immune response are part of the epigenetic network which contribute to tumour progression and metastatic evolution of the primary tumour cells (Coomber *et al.*, 2003).

1.4.2 Role of angiogenesis in metastatic evolution

One of the crucial elements of the metastatic pathway is the formation of new blood vessels, or angiogenesis, that enables tumours to escape from their primary site and enter into the blood circulation (Zetter, 1998). Each stage in the metastatic cascades *i.e.* local invasion, intravasation, survival in the vasculature, extravasation and eventually

colonization of the metastasising cells is related to angiogenesis. Various factors contribute to angiogenesis, notably the overexpression of vascular endothelial growth factor (VEGF) (Ellis & Fidler, 1996). Endothelial cell migration is an essential step in angiogenesis which is further governed by actin, its major cytoskeleton component, which can be remodeled into different structures such as filopodia (long filamentous structures), lammellipodia (cytoplasmic protrusions) and stress fibers (actin filaments with inverted polarity) contributing to endothelial cell motility (Theriot & Mitchison, 1991; Disanza *et al.*, 2005). In this regard, endothelial cell migration is characterised by chemotaxis mediated by VEGF and basic fibroblast growth factor (bFGF), directional migration towards ligand (haptotaxis) activated by integrins binding to ECM and migration due to shear mechanical forces (mechanotaxis) mediated by adhesive forces (Senger *et al.*, 1996; Li *et al.*, 2002). VEGF is a major potent angiogenic factor that promotes proliferation and migration of endothelial cells (Rousseau *et al.*, 2000). Furthermore, VEGF binds to membrane tyrosine kinase receptors VEFGRs (VEGFR1, VEFGR2 and VEGFR3) mediating to downstream signals migratory pathways of endothelial cells (Olsson *et al.*, 2006). In addition, autophosphorylation and activation of p38 pathways contributes to lammellipodia formation (Lamalice *et al.*, 2004) whereas activation of PI3K by VEGFR2 further activates AKT/PKB (protein kinase B) contributing to EC motility promoting angiogenesis (Bernatchez *et al.*, 2001). Likewise, VEGF promotes reactive oxygen species (ROS) production by activating NADPH oxidase through activation of VEGFR2/PIK3/AKT/PKB that further promotes angiogenesis (Dimmeler *et al.*, 2000; Ushio-Fukai *et al.*, 2002). In addition, tumour suppressor genes, oncogenes and factors that contribute to epigenetic networks such as hypoxia, hypoglycemia and cytokine stimulation are known to influence angiogenesis in tumour metastases (Coomber *et al.*, 2003).

Moreover, formation of new blood vessels for nutrient supply by tumours cells on their own is regarded as a hallmark of cancer (Hanahan & Weinberg, 2000). An experiment carried out by Weidner *et al* (Weidner *et al.*, 1991) using endothelial cells from invasive breast carcinoma patients to examine the formation of blood vessels showed that numbers of and density of blood vessels correlates with the outcome of metastatic disease. An *in vivo* study of angiogenesis in gyanaecological leiomysarcoma has shown that tumour cells capable of disseminating into a chick embryo showed a strong angiogenic response (Alias *et al.*, 2015) providing further evidence of angiogenesis in the metastatic potential of primary tumours regardless of their site of origin.

1.4.3 Role of tumour hypoxia in invasion and metastasis

Tumour cells often adapt themselves to hypoxia, which is associated with metastasis (Mendoza & Khanna, 2009). Hypoxia is known to induce various events in tumours that contribute to angiogenesis, invasion and metastasis increasing the malignancy of tumours (Zhong *et al.*, 1999; Coomber *et al.*, 2003)A hypoxic microenvironment is associated with two of the very prominent characteristics of solid tumours, glycolysis and angiogenesis which promotes metastasis and invasion in tumours (Zhong *et al.*, 1999). The *HIF1α* subunit regulates the transcriptional activity of *HIF1*, which activates *VEGF* (resulting in an angiogenic response) and genes involved in glucose transporter and glycolytic enzymes (altering the metabolic nature of the tumour cells) (Webster, 1987; Lin *et al.*, 2004a). Previous studies (Zhong *et al.*, 1999; Mendoza & Khanna, 2009) have shown that *HIF1α* expression was higher in metastatic breast tumours (69%) compared to its expression in primary breast tumours (29%) which suggests that the increased expression of *HIF1α* contributes to cancer metastases (Zhong *et al.*, 1999).

1.4.4 Tumour microenvironment

One of the key determinants of the fate of metastasising tumour cells is the tumour microenvironment, which can change the behaviour and gene expression patterns of the cancer cells influencing their ability to grow and proliferate. The gene expression patterns, level of various proteolytic enzymes expressed and resistance to chemotherapy may differ if the same cancer cells are grown in two different microenvironments (Nakajima *et al.*, 1990). In addition, a microenvironment may be more selective to a type of tumour cells and may be more supportive to other types suggesting the preference of different tumour cells to different microenvironments influencing different metastatic outcome (Kuo *et al.*, 1995). The status of *p53* in cancer cells may be a determinant for anti-angiogenic therapy and growth whereas the expression of soluble cytokines like *IL-6*, and adhesion molecules could be a contributor for microenvironment mediated chemotherapeutic and apoptotic response (Lowe *et al.*, 2014). Similarly, the metastatic ability of the tumour cells in secondary organs are supressed by metastatic suppressor genes, which do not in general, contribute to regulation or suppression of primary tumours. Likewise, expression of chemokine receptors by tumour cells plays a role in supporting lymphocytes and other tumour cells in specific organs suggesting a role in organ specific metastases, in which both the tumour microenvironment and the metastasising tumour cells express the identical chemokine receptors (Müller *et al.*, 2001). For instance, experimental models in animals has shown that breast cancer cells with increased levels of *CXCR4* and *CCR7* metastasise to the lungs, breasts, liver and bone were increased levels of *CXCR12* and *CCL21* (Müller *et al.*, 2001). In contrast, inhibition of *CXCR4* expression is found to suppress breast cancer metastasis in these sites (Burger & Kipps, 2006). Similarly,

activated *RAS* mutations favours the metastatic growth and is associated with maintaining micrometastases and increased growth potential of the metastasised cells in secondary sites (Nomoto *et al.*, 1998).

1.4.5 Epithelial-Mesenchymal Transition (EMT)

An important group of molecules contributing to the process of metastasis are transcription factors such as *Slug*, *Snail*, *Twist* and *ZEB1* which are highly expressed in metastatic cells and play important roles in EMT (Beltran *et al.*, 2008; Talmadge & Fidler, 2010; Taube *et al.*, 2010). EMT, a concept developed from embryonic development, is defined as the reversible changes in a cellular phenotype resulting in a loss of epithelial characteristics and a gain of mesenchymal characteristics (Thiery & Sleeman, 2006; Kalluri & Weinberg, 2009) which includes stem cell properties (Mani *et al.*, 2008; Polyak & Weinberg, 2009) (Figure 1.6). Loss of adhesion and polarity of the epithelial phenotype and gain of increased motility through ECM of the mesenchymal phenotype contributes to migration and invasion that lead to increased metastases (Talmadge & Fidler, 2010). In addition, the process of EMT is associated with reduced expression of epithelial markers such as E-cadherin, cytokeratin and claudins leading to increased expression of mesenchymal markers such as N-cadherin and vimentin (Cates *et al.*, 2008). Various lines of evidence have been reported regarding involvement of EMT in tumour progression leading to invasion and metastases. (Acevedo *et al.*, 2007) reported that the overexpression of fibroblast growth factor receptor 1 in a mouse led to a tumour progression into adenocarcinoma through EMT. Similarly, a microarray based histochemical study of 479 invasive breast carcinoma samples showed the overexpression of mesenchymal markers and reduced expression of epithelial markers

in basal like subtypes (Sarrio *et al.*, 2008). In addition, micrometastatic cells from the bone of a breast cancer showed expression of mesenchymal phenotype with reduction of epithelial phenotype, which was also detected in primary breast tumours with aggressive phenotype *in vivo* (Willipinski-Stapelfeldt *et al.*, 2005). Moreover, EMT resulted from aberration in DNA methylation and histone modifications at genes associated with alterations in the signalling components and pathways such as reduced autophosphorylation of EGFR, increased oxidative phosphorylation, induced tumour cell chemotaxis and increasing survival signalling (Thomson *et al.*, 2011). In addition, genome-wide epigenetic reprogramming such as alterations in histone modification mark is identified in EMT driven chemoresistance and migration in mouse hepatocytes (McDonald *et al.*, 2011). Likewise, epigenetic and genetic dysregulation of microRNAs are reported to contribute to or suppress the process of EMT (Davalos *et al.*, 2012).

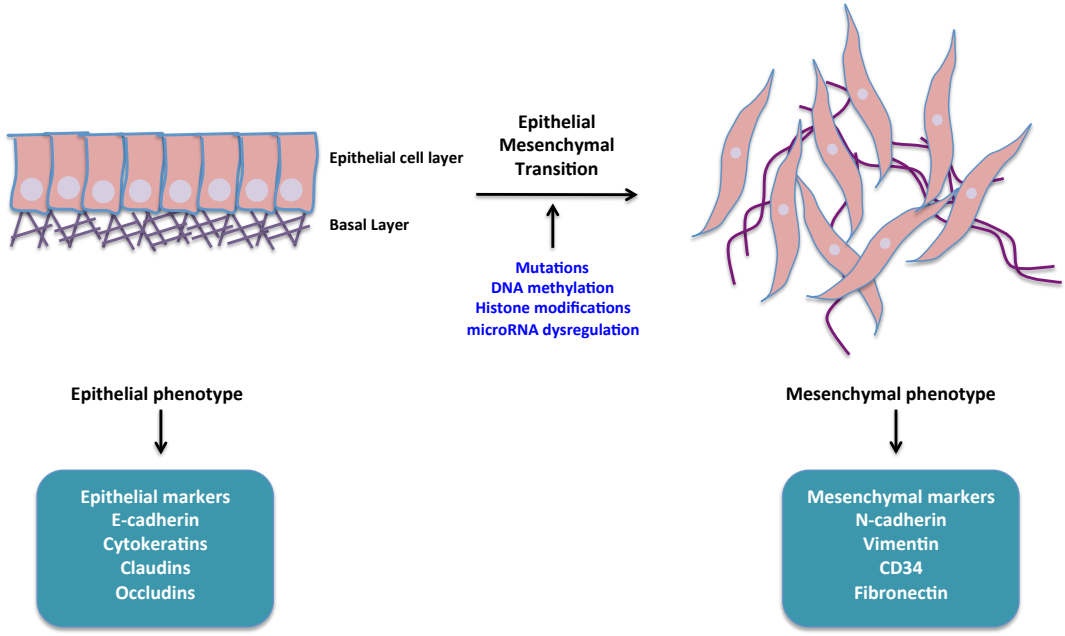


Figure 1.6: A process of Epithelial mesenchymal transition (EMT) in which epithelial cells characterised by adhesion, basal layer and lack of motility may be transformed into mesenchymal cells characterised by motility, lack of adhesion and degradation of basal layer or ECM resulting in an increase in migratory, invasive and metastatic behaviour and gain of stem cell properties. A number of factors such as mutation, DNA methylation, histone modification (such as alterations in histone methylation marks) and microRNA dysregulation contributes to EMT that results in change in epithelial markers to mesenchymal markers.

1.5 The Brain metastases

1.5.1 Patterns of brain metastases

Brain metastases, like any other tumour metastases, are not an exception to the seed and soil hypothesis, which has become a theoretical basis of cancer metastases (Zhang & Yu, 2011). Brain metastases account for up to 40% of all secondary tumours, with an estimated 27,000 new cases every year in the United Kingdom (Soffietti *et al.*, 2002; Gavrilovic & Posner, 2005). Current estimates suggest that 18-30% of patients with breast cancer eventually develop brain metastases (Weil *et al.*, 2005; Gori *et al.*, 2007; Kennecke *et al.*, 2007; Tosoni *et al.*, 2008). The prognosis of the brain tumours is very poor and both morbidity and mortality are high; the median survival of patients with brain tumour is 2 months without treatment and 12-15 with treatment (Subramani *et al.*, 2013), and overall the approximate median survival is 7 months (Sperduto *et al.*, 2010). The treatment for brain metastases generally includes whole brain radiotherapy (WBRT), surgery and chemotherapy (Zhang & Yu, 2011).

The underlying mechanisms of this organotropism towards specific secondary sites such as the brain are still poorly understood. Although the genetic basis of primary tumour formation is becoming increasingly clear (Stratton *et al.*, 2009) it is still unclear which of the many hundreds of tumour-associated alterations found in primary breast cancer (Curtis *et al.*, 2012; TCGA, 2012) contribute to metastasis and moreover, metastasis to specific secondary sites such as the brain.

One of the challenging hurdles in treating brain metastases is the failure of chemotherapy to cross through the tight junctions in blood brain barrier (BBB)

containing brain endothelial cells, arranged in a basement membrane including other supporting cells such as pericytes and astrocytes (Sugiyama *et al.*, 1999). In this regard, one of the intriguing questions is the period of latency between the presence of early circulating cells in the blood and appearance of well-developed brain metastases as a result of an ability of circulating tumour cells to cross the BBB (Riethdorf *et al.*, 2007). However, due to the leaky nature of the BBB, the drug resistance, to some extent, is attributed to expression of the membrane protein, P-glycoprotein which serves as a drug efflux pump in tight junctions. Moreover, the extravasated tumour cells, which are surrounded by activated astrocytes in the brain microenvironment are resistant to chemotherapy (Subramani *et al.*, 2013)

1.5.2 The brain microenvironment

The brain microenvironment (BME) consists of different types of neurons and glial cells such as astrocytes, microglia and other similar cells (oligodendroglia and satellite cells) (Zhang & Yu, 2011). In this regards, the BME is highly heterogeneous containing cells with drastically diverse biological functions maintaining homeostasis (Teschemacher *et al.*, 2015). The endothelial cells are present in the tight junctions in BBB that provides barrier to circulating molecules to enter into the brain, astrocytes provide nutrients maintaining homeostasis whereas microglia are resident macrophages that play roles as phagocytic cells in the BME (Wei *et al.*, 2013). Brain ventricles that separates the brain parenchyma (meninges) consists of Cerebrospinal fluid (CSF) maintaining a buffer system (Zhang & Yu, 2011). Hence, the BME represent the leptomeningeal space and a major site for the tumours extravasated to the brain, which is separated from the blood circulation by the tight junctions containing endothelial cells.

These prevents the entry of circulating molecules that otherwise would be allowed to pass through the blood vessel wall (Zhang & Yu, 2011). The presynaptic neurons release different neurotransmitters which activate the receptors on the postsynaptic neurons which require endogenous growth factors called neurotrophins such as nerve growth factors, brain-derived neurotrophic factors (BDGF), NT-3, NT-4 (Termini *et al.*, 2014). Interestingly, upregulation of neurotransmitters such as γ-amino butyric acid (GABA) has been reported in brain metastases, the GABA is catabolised by metastatic tumour cells through the GABA shunt to utilize it to form nicotinamide adenine dinucleotide (NADH). This suggests that the neurotransmitters may provide proliferative advantage to the metastatic tumour cells in the brain (Neman *et al.*, 2014). The tight junction in the BBB presents an obstruction for treatment of brain metastases and any tumours growing in the BME (Nakagawa *et al.*, 2009). Moreover, BBB may be leaky in the brain in comparison to the blood tumour barriers, which could be a reason for being resistant to chemotherapy (Fidler *et al.*, 2010). It has been suggested that a glioblastoma stem like cell may differentiate into endothelial cells to promote blood supply to the tumours. Moreover, the expression of VEGF and its regulation by integrins in activated tumour cells favours the tumour outgrowths in the brain microenvironment (Lorger *et al.*, 2009). The brain is also characterised by an ample supply of nutrients and oxygen and has a highest rate of glucose and oxygen metabolism, which is possible due to dense blood vessel networks. However, it has been observed that the cancer cells mainly depend on lactate production in aerobic conditions as an alternative source of energy (Teschemacher *et al.*, 2015), suggesting an alternative way of deriving energy to survive and to proliferate.

1.5.3 Role of brain microenvironment in brain metastases

Brain metastases account for up to 40% of all solid tumours and are extremely difficult to treat (Lorger & Felding-Habermann, 2010; Subramani *et al.*, 2013). It has been revealed that the BME primarily consists of endothelial cells, microglia and astrocytes as its resident cells in which circulating tumour cells invading brain first interact with the endothelial cells at the BBB (Lorger & Felding-Habermann, 2010). The endothelial cells at the BBB that acts as the barrier for arresting tumour cells later supports its growth. However, the mechanism by which circulating tumour cells negotiates with the endothelial cells and invades through the BBB maintaining it intact still remains elusive. Similarly, the BME, that is highly selective for any foreign cellular entity, consists of reactive astrocytes and re-localizes to associate with the incoming tumour cells (Klein *et al.*, 2015). The astrocytes are mainly housekeeping cells to supply nutrients and maintain homeostasis in the BME by playing a key role in maintaining concentration of extracellular glutamate, ions, pH as well as providing metabolic substrates for neurons (Marchetti *et al.*, 2000). Moreover, any foreign or tumour cells stimulate astrocytes to become reactive, which are otherwise non proliferating and quiescent cells in the brain (Termini *et al.*, 2014). It has been shown that the reactive astrocytes secrete cytokines, heparanase, neurotrophic factors such as transforming growth factors, stromal cell derived factor 1 (SDF-1), sphingosen-1 phosphate and further activate MMP9, which promotes the growth of the tumour cells (Hoelzinger *et al.*, 2007). MMP9 is a pro-angiogenic factor that promotes tumour cell proliferation and invasion by secreting growth factors from the extracellular matrix (Du *et al.*, 2008). Similarly, microglia are phagocytic cells and are resident macrophages capable of presenting antigen, which are overcome by tumour cells by production of immunosuppressive factors (Murata *et al.*,

1997). Microglia stimulates angiogenesis, tumour proliferation and invasion by promoting immune suppression and secretion of various growth factors, cytokines and enzymes (Lorger & Felding-Habermann, 2010). However, the role of microglia in exhibiting cytotoxic effects in lung cancer metastases in the brain suggests its diverse and varying roles in different stages and heterogeneity of the brain metastases (Lorger & Felding-Habermann, 2010). Subramani, *et. al.* (Subramani *et al.*, 2013) co-cultured astrocytes with melanoma, breast and lung tumour cells to investigate a unique interaction of tumour cells with the brain microenvironment in brain metastases and observed that the astrocytes promotes survival and chemo-protection to the tumour cells. Likewise, it is possible that the tumour cells which have an ability to utilize nutrients and signalling cascades to overcome the natural barrier are selected by the BME (Termini *et al.*, 2014). Furthermore, overexpression of *BMP-2* by metastatic cells promotes neural stem cells (NSCs) to differentiate into astrocytes, which in turn promotes tumour survival as a micrometastases or as dormant cells by cytokine stimulation (Neman *et al.*, 2013). Initial survival of these metastatic cells with support from astrocytes allow them to acquire more genetic and epigenetic dysregulation to develop neural phenotypes giving them selective advantages to further growth (Subramani *et al.*, 2013; Termini *et al.*, 2014).

1.5.4 Breast to brain metastases (BBM)

CNS metastases from breast tumours comprise of 78% multiple brain metastases, 14% solitary brain metastases and 8% leptomeningeal brain metastases (Lin *et al.*, 2004b). Patients' younger age, premenopausal status, ER-, PR – and HER2- (triple negative) status are associated with increased frequency of CNS metastases (Tham *et al.*, 2006).

The nodal HER2 status, tumour size and high tumour grade are also associated with CNS metastases (Pestalozzi *et al.*, 2006).

Experimental evidence indicating that tumours originating in specific organs favour certain sites of metastasis has existed for over fifty years (Sugarbaker, 1952) However, the underlying mechanisms of this organotropism towards specific secondary sites such as the brain are still poorly understood. Although the genetic basis of primary tumour formation is becoming increasingly clear (Stratton *et al.*, 2009), it is still unclear which of the many hundreds of tumour-associated alterations found in primary breast cancer (Curtis *et al.*, 2012; TCGA, 2012) contributes to metastasis and moreover, metastasis to specific secondary sites such as the brain.

The common long lag-time between primary tumour diagnosis and recurrence of a detectable secondary tumour in ER+ve breast tumours suggests that cells from these tumours undergo a period of dormancy (Aguirre-Ghiso, 2007; Lim *et al.*, 2012). These dormant cells are often found as micrometastases in bone marrow. However the presence of these micrometastases is not in itself a strong prognostic indicator for later metastatic disease (Klein, 2003; Pantel *et al.*, 2008). Recent trials of long-term tamoxifen treatment suggests that ER signaling inhibition is capable of suppressing proliferation of dormant cancer cells but is not capable of killing them (Davies *et al.*, 2013; Zhang *et al.*, 2013).

Disseminated ER+ve tumour cells must acquire further adaptive changes to enable intracranial growth. Tumour cells that infiltrate the brain may enter a state of micrometastatic growth where the rate of proliferation is counterbalanced by cell death (Demicheli, 2001). During this period of metastatic latency the general instability of the tumour genome can result in the evolution of cells capable of proliferating within the

microenvironment of the brain parenchyma. Alternatively, this evolution may occur in a different organ, most likely bone marrow (Zhang *et al.*, 2013) and the resulting cells make a second migration to the brain.

Primary breast tumour cells after disseminating from their original niche evolve into metastases through two possible routes. They may reach the lung capillaries, through the heart to reach to the systemic circulation (Pestalozzi, 2009). Therefore, breast cancer cells need to invade the barrier of lung capillaries to reach the blood from where they will extravasate to the distant organs. Alternatively, it is possible that the breast tumour cells may reach a lymph node through the invasion of lymphatic vessel, where the tumour cells either grow into lymph node metastases or will enter into the blood vessel (Chambers *et al.*, 2002). The tumour cells either enter into the blood directly or through the newly formed blood vessel from which they will eventually colonise into the brain.

One of the most intriguing factors underlying BBM is an adaptation of neural properties by primary breast tumour cells after they extravasated to the brain (Neman *et al.*, 2014). This neural-like phenotypic change may be a necessary adaptive change driven by the selection pressure of the neural niche (Van Swearingen *et al.*, 2014). It has been shown that the human brain exerts higher concentration of GABA as a neurotrophic factor, which play roles in cellular proliferation, migration, neural differentiation and death, which is catabolized into succinate to release NADH to fulfill a nutritional requirement (Neman *et al.*, 2014; Termini *et al.*, 2014). Interestingly, GABA is reported to be an alternative metabolite in various primary cancers such as gastric, colon, ovarian and breasts tumours including BBM from HER2+ and triple negative breast tumour cells when the TCA cycle and normal cellular metabolism is not favorable to them (Neman *et al.*, 2014). In addition, various proteins involved in GABA metabolism such as GABA

transaminase, glutamate decarboxylase, GABA transporter, reelin and parvalbumin are expressed by BBM tumour cells (Van Swearingen *et al.*, 2014). Moreover, a subset of GABAergic neurons expresses the protein reelin, which in association with HER2 is involved in motility and maintaining cytoskeleton by activating PI3K, which further elicits AKT and other downstream signaling molecules resulting in actin skeleton activation (Chai *et al.*, 2009). Furthermore, ENAH/HMena, a regulator for actin cytoskeleton is overexpressed in metastatic HER2+ cells, this is stabilized in the brain microenvironment by reelin and HER2, which further helps to modify cytoskeleton to promote tumour-stromal interaction in the extracellular matrix (Neman *et al.*, 2014).

1.6 The biology of breast tumour

1.6.1 Primary breast tumour and its subtypes

Breast cancer, a leading cause of cancer death in less developed countries, is also a common malignancy and a second leading cause of cancer deaths in women in the western world (Cianfrocca & Goldstein, 2004; Torre *et al.*, 2015) with estimated global deaths of 521900 in 2012 alone (Torre *et al.*, 2015). It originates from normal epithelial cells, characterised either as precursor lesions called atypical ductal hyperplasia (ADH), flat epithelial atypia (FEA), and ductal carcinoma in situ (DCIS) or more invasive phenotype invasive ductal carcinoma (IDC) (Park *et al.*, 2011). DCIS is a pre invasive form of breast cancer characterised by a proliferation of epithelial layers without invading it (Moelans *et al.*, 2011). 20% of patients carry DCIS, which are in general curable except for a minority of patients who go on to develop invasive ductal carcinoma (IDC) (Moelans *et al.*, 2010). It has been observed that the high and low grade IDC is derived from the high and low grade DCIS respectively (Moelans *et al.*,

2010). Genetic profiling of breast tumours has revealed that the majority are derived from luminal epithelium or their progenitors suggesting that they are from an epithelial origin (Park *et al.*, 2011). High survival rates (~90%) of breast cancer patients following early detection suggests that this is crucial for effective treatment and survival of the patients (Tommasi *et al.*, 2009).

Other than the histological origin, breast tumours have been classified into subtypes based on gene expression profiling and genomic alterations such as DNA methylation in normal breast (epithelial) tissues and primary breast tumours. Five major subtypes of breast cancer have been reported based on gene expression studies *i.e.* basal like, HER2 enriched (ERBB2), luminal A, luminal B and normal breast like with distinct disease phenotype entities (Sorlie *et al.*, 2003). Exome sequencing of 103 patients with different subtypes has uncovered 4,985 candidate mutations in total, with frequent C to T transitions in CpG dinucleotides (Banerji *et al.*, 2012). Similarly, investigation of somatic mutations and copy number changes using exome sequencing in the genomes of 100 patients have further revealed mutations in 40 cancer genes including driver mutations in some genes not reported previously to be involved in breast cancer such as *AKT2*, *ARID1B*, *CASP8*, *CDKN1B*, *MAP3K1*, *MAP3K13*, *NCOR1*, *SMARCD1* and *TBX3*, which suggests an extreme genetic heterogeneity of primary breast tumours (Stephens *et al*., 2012). This study identified 7,241 somatic point mutations that included single base substitutions, missense, nonsense, indels (insertions and deletions) and frameshift mutations. Analyses of DNA methylation variable regions (MVRs) of estrogen receptor alpha (*ESR1*) gene in ER positive and negative cell lines have identified that the *ESR1* promoter was unmethylated and intragenic regions were methylated in ER+ cells, whereas *ESR1* promoter was methylated and intragenic

regions were unmethylated in ER- cells (Shenker *et al*., 2015). This revealed that there is a correlation between the ESR expression in intragenic methylation in homogenous cell populations further suggesting the significant differences between *in vivo* and heterogeneous tumour microenvironment. This further highlights the importance of looking in to the possibility of alterations in intragenic methylation in tumours contributing to epigenetic dysregulation of genes. Ronneberg *et al* (Ronneberg *et al.*, 2011) reported a cluster on myoepithelial origin and two clusters on luminal epithelial origin based on DNA methylation profiling. Interestingly, breast tumours subtypes with a similar genetic or gene expression profiling showed different epigenetic profiling. Histological and clinicopathological features of breast tumours samples have identified triple negative tumours as having poor prognosis and disease free survival (Haffty *et al.*, 2006; Onitilo *et al.*, 2009). Interestingly, BRCA mutations were commonly associated with triple negative tumours (Haffty *et al.*, 2006). Similarly, a previous study (Neve *et al.*, 2006) identified different patterns of copy number changes in breast cancer subtypes. Likewise, luminal B was found to have the highest percentage of lymph node metastases compared to other subtypes including luminal A (Inic *et al.*, 2014). HER2+ /ER- are associated with poor outcome due to the presence of residual tumours (Carey *et al.*, 2007) and ER+ tumours are noted to have metastatic dormancy for many years (Zhang *et al.*, 2013). An integrated genomic and transcriptomics profiling of 2000 breast cancer patients with a clinical follow up including copy number changes and SNPs identified various intermediate subgroups. ER+ subgroups had higher mortality risks and driver mutations, luminal A subgroups with low genomic instability. Luminal B subgroups showed frequent genomic translocations and basal like tumours had high genomic instability and good patient outcome (Curtis *et al*., 2012).

1.6.2 Evolution of the metastatic breast tumour

Breast cancer spread to other distant organs takes place in an organized fashion, first through the lymphatic system and then through the blood. The micrometastatic paradigm states that micrometastatic lesions may already exist in breast cancer patients at the time of primary breast cancer diagnosis and that these will later develop as metastatic disease (Cianfrocca & Goldstein, 2004). It has been shown that breast tumours may lose their tissue architecture due to aggressive and disorganized proliferation. Hence, inhibition of EGF signaling due to loss of 1-integrins and EGF may lead to the aberration of these structures (Aguirre-Ghiso, 2007). The tumour heterogeneity in terms of DNA methylation patterns in primary and metastatic breast tumours have identified aberrant DNA methylation of *BMP6, BRCA1* and *CDKN2A* in lymph node metastases that may evolve as a distant metastases in the future (Barekati *et al*., 2012). Moreover, analyses of single cell progenies from a population of metastatic breast cancer cells from an immunodeficient mouse has showen that the different cell progenies favor metastases to different distant organs suggesting a different requirement for metastases to different secondary sites (Minn *et al*., 2005b). Likewise, genomic hypomethylation and hypermethylation events in some important genes such as *CDH1*, *CST6*, *EGFR*, *SNAI2* and *ZEB2* are associated with EMT leading to metastasis to different secondary sites (Rodenhiser *et al.*, 2008). This suggests that breast tumours could possibly possess an inherent capacity to metastasise to different organs (Weigelt *et al.*, 2005).

1.6.3 Prognostic and predictive indicators for breast tumour recurrence

1.6.3.1 Tumour phenotype, patients' age/ethnicity and vascular/lymphatic status

The presence of breast tumour in axillary nodes *i.e*. tumour size and the number of lymph nodes involved are prognostic factors for distant recurrence of breast tumours. The number of lymph nodes and tumour size are independent to each other. However, both of these factors together increase the chances of distant metastases. Moreover, distant recurrence of breast tumours increases with the tumour size and the median time to develop distant metastases decreases as the tumour size increases (Carter & Allen, 1989). Similarly, tumour grade that ranges from 1 to 3 is calculated based on the total score for its mitotic index, differentiation and pleomorphism for which the score ranges from 1 to 3. Well differentiated tumours (grade 1) have scores from 1 to 3, moderately differentiated tumours have scores from 6-7 (grade 2) where poorly differentiated tumours have scores from 8-9 (grade 3). Tumours grades are more important especially for lymph node negative patients who have borderline tumour size (Cianfrocca & Goldstein, 2004). Furthermore, African American and Hispanic women have reduced survival rate in comparison to American women, this is attributed to several factors such as lack of clinical care (Daly *et al*., 1985). Breast cancer patients younger than 35 years of age generally show worse prognosis (Albain *et al*., 1994). The presence or absence of breast tumour in axillary lymph node is a very important prognostic indicator for early stage breast cancer, the risk of breast cancer recurrence is directly associated with the number of axillary lymph nodes involved. Based on the number of axillary lymph node involvement breast cancer patients are classified as negative lymph nodes, 1-3 lymph nodes, 4-9 lymph nodes and more than 10 lymph nodes. Long term follow up of breast cancer patients have shown a high risk of breast cancer related death or tumour recurrence due to breast cancer micrometastasis in lymph nodes (Saez *et al*., 1989).

Similarly, local and distance metastasis of breast cancer increases with vascular invasion of lymphatic vessel. Vascular invasion has been associated with recurrence risk and is more important in terms of decision-making in breast cancer patients with borderline tumours size and without lymph node metastasis (Cianfrocca & Goldstein, 2004).

1.6.3.3 Hormone receptor (ER/PR/HER2) Status

The presence of ER and PR is a crucial predictive factor in patients to assess the likelihood of benefitting from tamoxifen (an antagonist of estrogen receptor) treatment (Cianfrocca & Goldstein, 2004; Park *et al*., 2011). Statistically, ~10% of the patients are ER/PR/HER+, ~70% are ER/PR+, HER-, and ~13% are triple negative (Onitilo *et al*., 2009). This shows that, regardless of HER status, around 80% of the patients are ER/PR+. ER+ breast tumours can remain dormant or not detected for many years after the diagnosis of the primary tumour suggesting that the cancer cells may proliferate with a slow rate during the period of latency (Zhang *et al*., 2013). Moreover, early stage breast cancer may eventually metastasise after a long period of time, even decades after a diagnosis of primary tumours (Zhang *et al*., 2013). In contrast, recurrence of ER- breast tumours takes place within three years from the initial diagnosis and generally the recurrence rate decreases after five years. Therefore, ER+ breast tumours pose a substantial challenge (Hilsenbeck *et al*., 1998). Similarly, overexpression of the *HER2* gene denotes the aggressiveness of the diseases with reduced disease free and overall survival of breast cancer patients (Slamon *et al*., 2001). In patients that develop resistance to therapy the HDAC I inhibitor SNDX-275 (an epigenetic modulator) is observed to increase the ability of the patients to respond to herceptin therapy (Huang *et al*., 2011). Similarly, combination therapy using a HDAC II inhibitor with tamoxifen

has been effective in hormone therapy resistant patients (Munster *et al*., 2011).

1.6.4 Mechanisms of dormancy of disseminated breast tumour cells

Comparative Genomic Hybridization (CGH) in primary breast cancer has shown that the evolution of micrometastases or dormancy may depend on genetic dysregulation of genes in primary tumours leading to continued progression and systemic microevolution of tumours (Aguirre-Ghiso, 2007). Several components that regulate cancer cell viability and self-renewal mechanisms contribute to dormancy, this is probably controlled by cross talk between the dormant cells and their microenvironment (Zhang *et al*., 2013).

It is possible that the tumour cells, which remain dormant in a new niche, may be tumour initiating cells or cancer stem cells, which have the capacity to reinitiate the tumour. However, is not clear if the cells that reinitiate metastasis in bone are the same cells that contribute to tumorigenesis in primary tumours (Marsden *et al*., 2012). It has been shown that tumour initiating potential of breast micrometastases in the lung is suppressed by *BMP2* signaling. Furthermore, metastasis in lung is driven by *Coco* expression by terminating metastatic dormancy (Zhang *et al*., 2013). *WNT* and *Notch* pathways are activated to maintain tumour-initiating cells by the ECM protein TNC (Tenascin C).

1.6.5 Detection of micrometastases and late occurrences

Only micrometastases more than 300 μm in diameter can be detected by MRI spectroscopy (Zhang *et al*., 2013). Therefore, other techniques will be required to identify the majority of micrometastases. One promising technique to detect the

presence of micrometastases is the examination of tumour DNA in peripheral blood using suitable molecular methods (Weigelt *et al*., 2005). It is believed that the circulating tumour cells (CTCs) and tumour DNA in patients could signify the present of micrometastases in distant sites (Giuliano *et al*., 2011). Similarly, the presence of micrometastases could be examined by the presence of tumour free circulating DNA in the blood based on the DNA released by the micrometastases into the blood (Zhang *et al*., 2013). Histopathological features and gene expression profiles are used to predict early relapses (less than five years) with reasonable accuracy (Wang *et al*., 2005). Unlike late occurrences, early relapses may be a result of rapid proliferation of tumour cells. However, it has been extremely difficult to predict late occurrences, which generally are not correlated to rapid proliferation of the tumours (Cardoso *et al*., 2012).

A comparison of gene expression profiling of primary tumours and late reoccurring tumours has been used to identify gene expression signatures for late occurrences (Mittempergher *et al*., 2013). The results of this study suggest that there may be specific epigenetic and genetic changes that occur over a long evolutionary process where the distant site microenvironment provides a distinct selective pressure to that applied to the primary tumour. Therefore, it is important to examine primary tumours and lately recurred metastases to identify unique genetic and epigenetic dysregulation responsible for late recurrences of the primary tumours.

CHAPTER 2

Aims and Objectives

2.1 The Project Background

The aim of this project was to identify epigenetic alterations contributing to metastatic growth, which provides potential for tumour cell dissemination and infiltration to a new niche. We hypothesised that epigenetic silencing of BBM associated genes would either occur as

a) Early events, in which the epigenetic alterations occur early in the primary tumours that contribute to local invasion and intravasation**.** These early events may be required for specific distant site metastasis and also contributes to primary tumour development (Figure 2.1).
b) Late events, in which the epigenetic alterations occur late in BBM. These play no significant role in the initial evolution of the primary tumour but contribute to the development of the secondary brain metastasis, either by promoting invasion or improving the capacity for these foreign cells to survive in the new microenvironment of the brain (Figure 2.1).

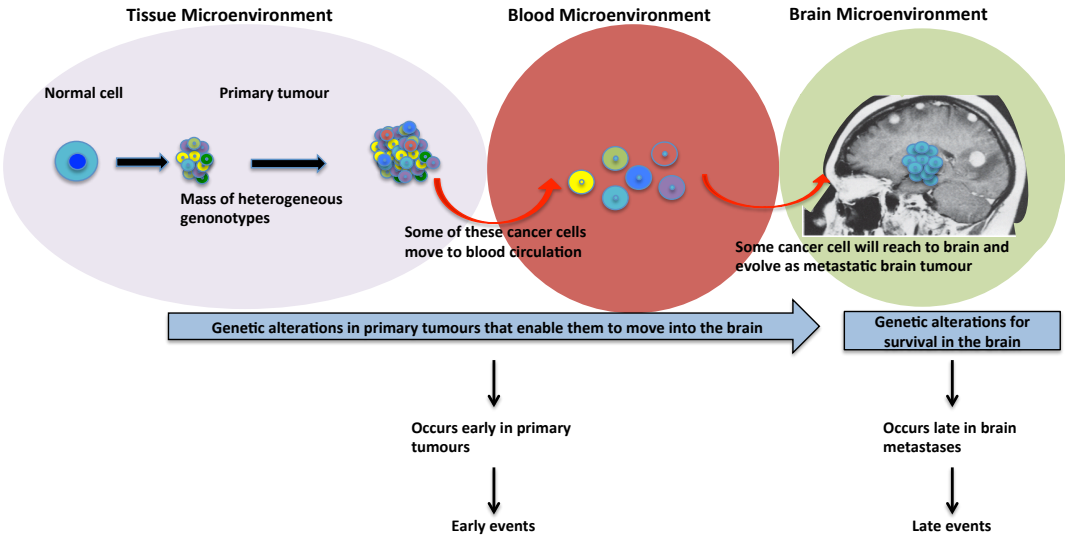


Figure 2.1: The possible genomic alterations in cancer metastasis to the brain during metastatic cascade. The genomic alterations that occur in primary tumours (early events) in a malignant phenotype may provide a selective advantage to tumour cells to invade the surrounding tissues and enter the blood vasculature, (intravasation), which eventually reaches the brain (extravasation). Similarly, additional genomic alterations that occur in the brain microenvironment (late events) may provide a selective advantage to metastasised tumours cells to grow and to evolve into the malignant brain metastases.

In this regard, the project intended to identify those genes which are dysregulated either early in primary tumours (early events) that provided these cells a selective advantage to metastasise to the brain or late in the metastasised cells in the brain (late events) that provided potential to those cells survive against the selective pressure in order to grow and to proliferate in the brain microenvironment.

2.2 Identifying genes dysregulated in BBM *i.e.* genes frequently methylated in brain metastases.

The first step was to identify genes dysregulated in BBM. This was achieved as follows.

2.2.1 A broad literature search to screen genes frequently methylated in one of the tumour types that metastasises to the brain

Given that lung, breast, melanoma, colorectal and renal cancers are the most frequent ones metastasising to the brain, there may be some genomic alterations common between them (chapter 4). There may be a possibility that at least in part, genomic alterations that drive tumour formation in these primary organs provide the potential for colonization of a distinct subset of primary tumours in secondary sites. In order to choose candidate genes that are uniquely dysregulated in BBM but not in primary breast, the initial screenings included only those genes, which are either methylated in lung (figure 2.2, gene B), melanoma (Figure 2.2, gene C), or RCC (Figure 2.2 gene D).

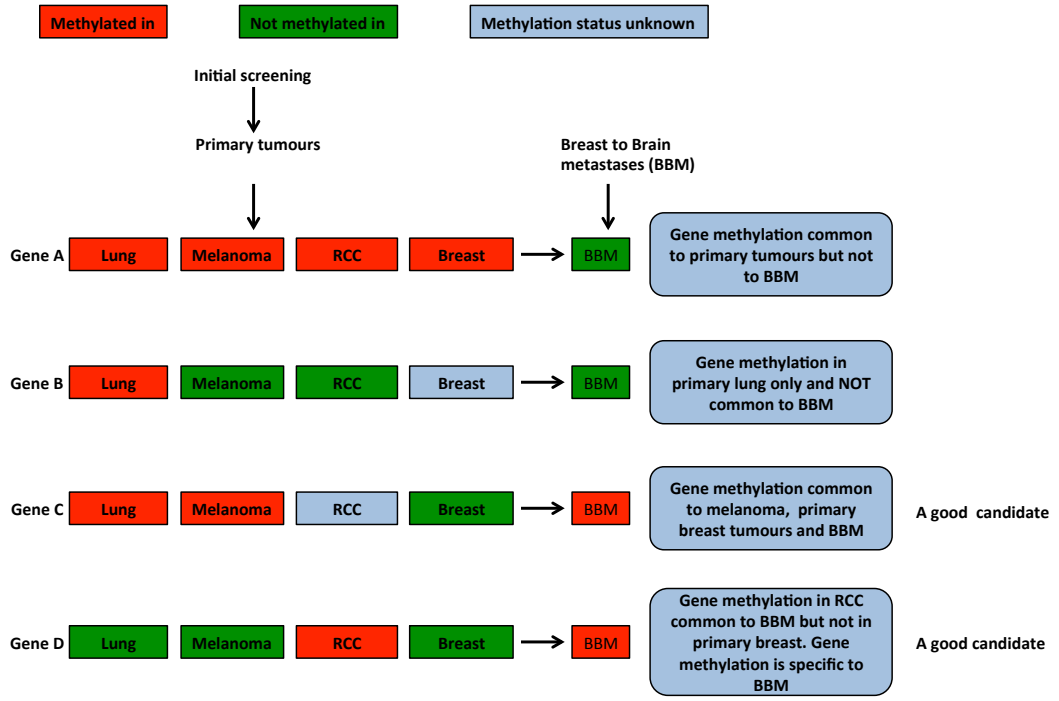


Figure 2.2: Screen of dysregulated genes in BBM. The candidate genes methylated in all primary tumours that readily metastases to the brain contribute to primary tumours including breast (*Gene A*). Initial screening included the genes, which are methylated either in lung (*Gene B*), melanoma (*Gene C*) or RCC (*Gene D*) but not in primary breast. However, some genes with their unknown methylation status in primary breast (such as *Gene B*) were also chosen. A good candidate would be methylated in one of the tumour types that readily metastasises to the brain such as lungs, melanoma and RCC but not methylated in breast (from the initial literature search), and frequently methylated in BBM samples identified by laboratory analyses (such as *gene C, D*).

2.2.2 A broad literature search for screening of genes downregulated in EMT

The process of EMT is involved in metastasis and invasion (section 1.4.5). We hypothesise that the genes that are downregulated in the process of EMT may provide a selective advantage to cells to attain mesenchymal phenotypes increasing their motility suitable for invasion and metastases. In order to extend the list of candidate genes dysregulated in BBM, those genes downregulated in EMT process were selected. The genes downregulated in EMT included occludins, claudins, keratins and tumour suppressor microRNAs (chapter 4).

2.2.3 Bioinformatic analysis of The Cancer Genome Atlas (TCGA) data

We hypothesized that gene methylation that may contribute to BBM will occur commonly in primary lung tumours, as these often metastasise to the brain in a short time period relative to initial diagnosis. These metastasising breast tumours may have genomic alterations common to lung tumours that are responsible for BBM. Therefore, the genes that are infrequently methylated in non-metastasising breast tumours and frequently methylated in primary lung tumours (that readily metastasis to the brain) may be found to be commonly methylated in metastatic brain tumours that derive from both lung and breast tumours. In order to identify such novel candidate genes that contributes to BBM, analyses of genome wide 450K methylation array data was carried out.
(chapter 5, figure 2.3)

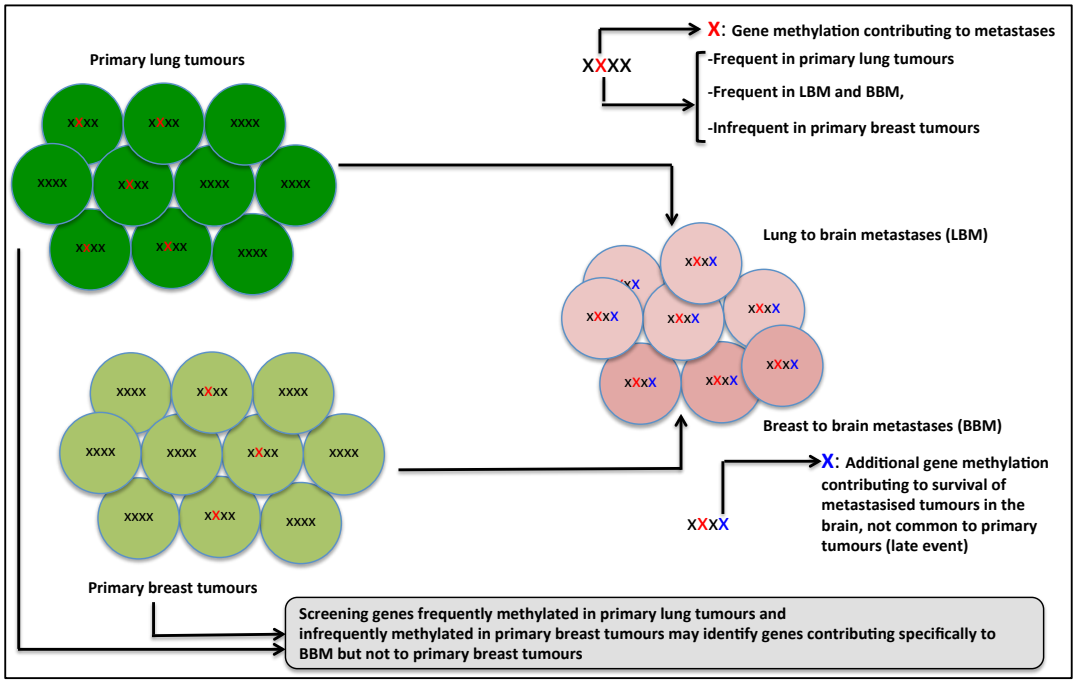


Figure 2.3: The bioinformatics analysis of the Infinium BeadChip 450K methylation array data for primary breast and lung tumours from TCGA. The frequent genomic dysregulation (X) in a genome (XXXX) responsible for primary lung tumours may also be common and frequent in lung to brain metastases brain metastases (LBM) and BBM, which, however, may be infrequent but common in primary breast tumours. Additional genomic dysregulation (X) may occur in the genome of BBM tumour (XXXX) inside the brain microenvironment that may contribute to survival of brain micrometastases and to the eventual development of macrometastases.

2.3 Investigation of methylation status of candidate genes and their expression primary breast tumours and BBM

The methylation status of candidate genes obtained from a literature search and TCGA data analysis were investigated using Combined bisulphite and restriction analysis (CoBRA) in BBM samples. The candidate genes frequently methylated in BBM were further investigated in an independent cohort of primary breast tumours. Furthermore, those candidate genes which were differentially methylated in primary breast tumours and BBM samples were analysed for their expression in BBM samples to ensure that the genes were silenced due to promoter hypermethylation (chapter 4 and 5). The summary of the methodology is illustrated in figure 2.4.

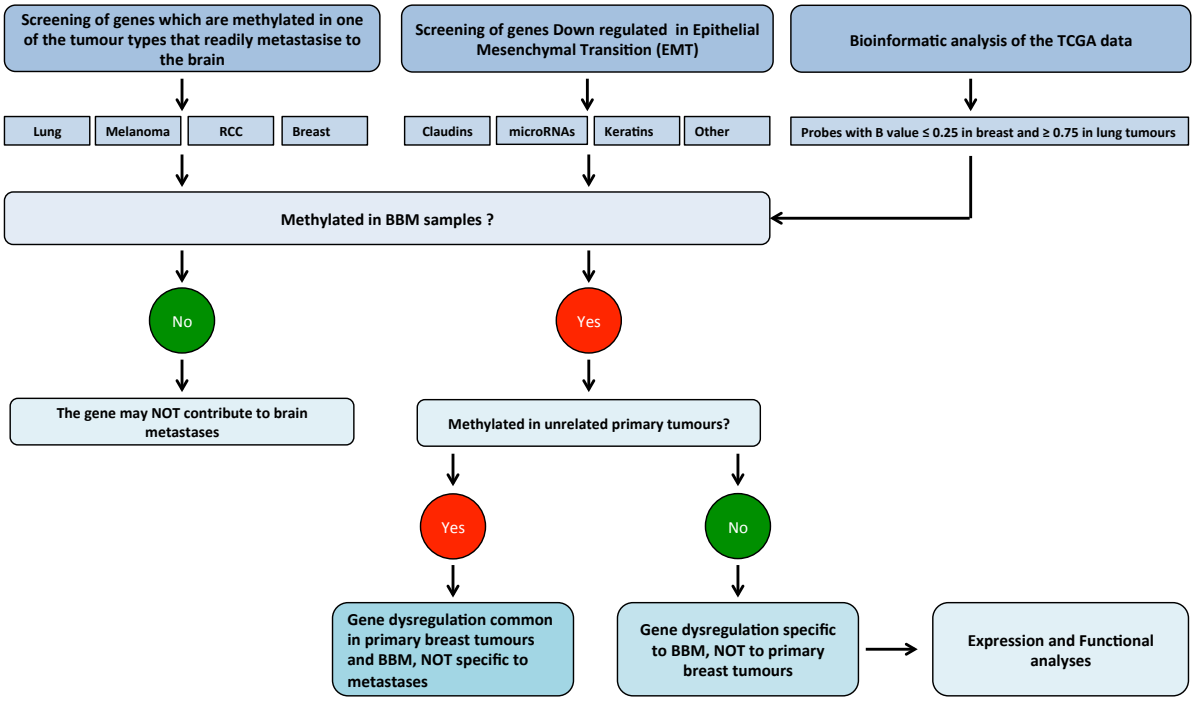


Figure 2.4: Graphical overview of methodologies used in screening and experimental validation of genes dysregulated in BBM. The candidate metastatic suppressor genes screened from a literature review to identify those genes that are methylated in lung,

melanoma and renal cancer as these often metastasise to the brain rapidly. If these genes were not known to be frequently methylated in breast tumours (that metastasise to the brain with a longer lag period), they were considered as good candidates. Similarly, a literature review included candidate genes downregulated in EMT. In addition, bioinformatic analyses of Genome wide 450K methylation data for primary lung and breast tumours from the TCGA were carried out to identify novel candidate genes. The identified genes were further screened in BBM samples using CoBRA and those that were frequently methylated in BBM samples were screened for methylation in primary breast tumours. Those that were frequently methylated in BBM samples and infrequently methylated in primary breast tumours were analysed for their expression in BBM samples and other functional roles.

2.4. Study of metastatic potential of candidate genes silenced or downregulated in BBM due to promoter hypermethylation

To achieve this aim, two independent sets of experiments were carried out.

2.4.1 Wound healing assay

In a wound-healing assay the breast cancer cell lines knocked down with siRNA nucleotide oligos against a respective gene being studied were seeded as a monolayer in 6 well plates (figure 2.5A). The wound/gap was created using 200µl pipette tip (section 3.6.3). The migratory potential of the cells was determined based on the distance travelled by cells towards the gap using statistical measurement (chapter 6).

2.4.2 Invasion assay

In a transwell assay, the breast cancer cell lines knocked down with siRNA oligos against respective gene were seeded as serum free media in a upper surface of a trans well contained in 6 well plates (figure 2.5B). The transwell experiment was prepared by coating an 8µm pore basement membrane with matrigel (section 3.6.4). The invasive potential of the cell lines were determined based on the number of cells invaded through the matrigel coated base membrane (section 6).

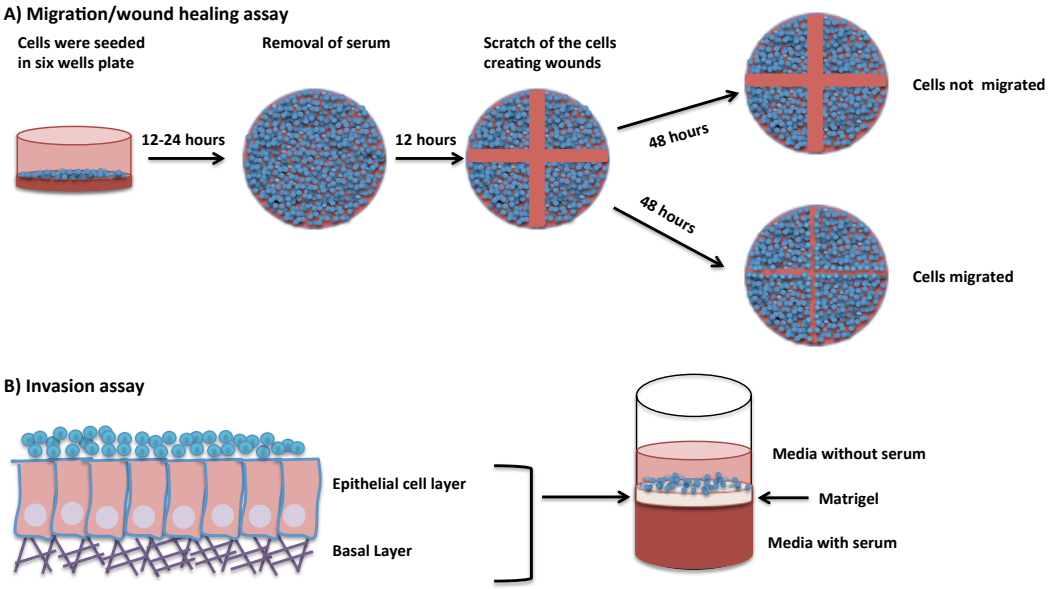


Figure 2.5: The experiments to investigate metastatic potential of breast cancer cell following the knockdown of candidate genes. The breast cancer cell lines knocked down with siRNA oligos against the candidate gene were used in a wound healing assay (A) and a trans well invasion assay (B) to investigate migratory and invasive potential of breast cancer cell lines respectively.

2.5 Identification of novel candidate genes by genome wide analyses of aberration in DNA methylation contributing to BBM

This aim was achieved by Illumina BeadChip Genome Wide 450K methylation analysis using 24 BBM samples (section 3.7.1). This analysis identified novel candidate genes, which are either hypomethylated or hypermethylated in BBM compared to primary breast tumours from the patients (chapter 7).

2.6 Investigate whether the genomic dysregulation in BBM is an early or a late event during the evolution of BBM in individual patients

The genetic dysregulation in the process of metastases is either an early or a late event (section 2.1, figure 2.1). This aim was achieved by determining the methylation status of candidate genes in primary tumours and their corresponding BBM samples in individual patients (chapter 4, 5 and 7).

2.7 Investigate whether the candidate genes identified (metastatic suppressor or metastatic promoter) genes could be used as prognostic biomarkers for BBM

To achieve this aim the methylation status of a panel of potential candidate genes (prognostic markers) was determined using Methylation Specific PCR (MSP) using tumour-free circulating DNA (section 3.3.4.2) isolated from individual patients' plasma (section 3.2.2) taken at the time of BBM surgery. The methylation status of these genes in plasma samples and BBM samples was used to determine if this panel of genescould potentially be used as prognostic biomarkers for BBM (chapter 7).

CHAPTER 3

Materials and Methods

3.1 Patients and Samples

3.1.1 Breast to brain metastases (BBM) samples

The Walton Research Tissue Bank (WRTB), Liverpool and Brain Tumour North West (BTNW), Preston provided fresh frozen metastatic brain tumour samples that had originated from breast, lung, melanoma, Renal Cell Carcinoma (RCC) and colorectal cancer. Genomic DNA was extracted from all the tumours collected. The metastatic brain tumour samples from breast, lung, melanoma and RCC were designated as BBM, LBM, MBM and RBM respectively. A list of brain metastases samples is given in Appendix A1.

3.1.2 Matched primary breast tumours

Twelve Formalin-Fixed Paraffin Embedded (FFPE) matched primary breast tumours corresponding to metastatic brain tumours were provided by BTNW tissue bank, Preston. The list of FFPE matched primary tumours and normal breast is given in table 3.1.

National research ethics committee approved tissues from the research banks and informed consent was obtained from each patient. This study was conducted according to the principles expressed in the Declaration of Helsinki.

Breast to brain metastases (BBM) code	Primary breast tumour code	Normal breast tissue
BM1	BP1	BN1
BM2	BP2	BN2
BM3	BP3	
BM5	BP5	
BM8	BP8	BN8
BM10	BP10	
BM11	BP11	
BM12	BP12	
BM13	BP13	
BM14	BP14	
BM15	BP15	

Table 3.1: The matched pairs of tumours; BBM, primary breast tumours (BP) and normal breast tissues (BN) from individual patients. Eleven primary tumours and only three normal breast tissues were available from individual BBM patients.

3.1.3 Genomic DNA from primary breast tumour

A cohort of 60 independent primary breast tumours were provided by Ivan Bièche (Department of Genetics, Institute Curie, Paris) and Farida Latif (Department of medical genetics, University of Birmingham). The primary breast tumours are denoted as BP for convenience. The list of primary tumours is given in appendix A2.

3.1.4 Plasma/serum of patients from BBM

Ten plasma and serum samples, collected at the time of surgery of BBM, were provided by BTNW. A list of Plasma/Serum Samples collected is given in table 3.2.

Local research committees approved collection of samples and informed consent was obtained from each patient. This study was conducted according to the principles expressed in the Declaration of Helsinki.

Breast to brain metastases (BBM) code	patient serum code
BM1	SD1
BM2	SD2
BM5	SD5
BM6	SD6
BM7	SD7
BM8	SD8
BM10	SD10
BM11	SD11
BM12	SD12
BM13	SD13

Table 3.2: A cohort of patients' serum and their corresponding breast to brain metastases samples from individual patients. These serum samples were collected at the time of brain tumour surgery from the patients, which were used to carry out MSP to find if the methylation status of the BBM samples correlates with the serum samples (chapter 7).

3.1.5 Breast cancer cell lines

Five breast cancer cell lines (MCF7, T74D, MDA-MB231, BT549 and ZR75) were used for tissues culture purposes and to carry out experiments. Prof. Weiguang Wang at the Research Institute of Healthcare Sciences (RIHS), University of Wolverhampton provided these cell lines. They were routinely maintained in DMEM (Sigma, UK) supplemented with 10% FCS at 37°C, and 5% CO_2.

3.2 Nucleic Acid Extractions and Preparation

3.2.1 Genomic DNA Extraction

Genomic DNA was extracted from fresh-frozen metastatic brain tumours, FFPE samples for matched primary tumours and also from patients' plasma taken at the time of metastatic brain tumour surgery.

3.2.1.1 Genomic DNA extraction from fresh frozen metastatic brain tumours

Genomic DNA was extracted from fresh-frozen metastatic brain tumours using *DNA isolation kit from cells and tissues* (Roche, Germany). 25mg of fresh frozen tissue was ground with a mortar and pestle in dry ice and 500µl of lysis buffer was added immediately to lyse the cells. The samples were homogenized using a syringe until the fine suspension was obtained and incubated at 37°C for 30 minutes. 1µl of Proteinase K solution was added to each sample, vortexed to mix, and incubated at 65°C in order to degrade proteins. 25µl of RNase solution was added to each sample, vortexed to mix and was incubated at 37°C for 30 minutes to ensure the degradation of RNA. 250µl of protein precipitation solution was added to each sample, vortexed thoroughly for 5-10 seconds, placed on ice for 5 minutes and each sample was centrifuged at 17000g for 30 minutes at room temperature. Supernatant containing DNA was transferred carefully to a new tube, 0.7 volumes of isopropanol was added to the sample to precipitate followed by centrifugation at 1370g for 10 minutes. Precipitated gDNA was washed with 1.5 ml of ice cold 70% ethanol, centrifuged at 1370g for 5 minutes and the supernatant was discarded. The DNA pellet was air-dried until the ethanol was fully evaporated. DNA

was dissolved in an appropriate amount of water (50-100 µl) and was quantified using a *Nanodrop 2000* spectrophotometer (Thermo scientific, USA).

3.2.1.2 Genomic DNA extraction from Formalin fixed FFPE tumour samples

Genomic DNA from FFPE samples was extracted using *QIAamp DNA FFPE tissue kit* (Qiagen, USA). Small blocks of samples embedded with paraffin were cut into eight sections (5 –10 µm thick) and excess paraffin was removed using a scalpel. These sections were immediately placed in a sterile 1.5 ml microcentrifuge tube, mixed with 1 ml xylene, vortexed vigorously for 10 seconds in order to dissolve paraffin and centrifuged at 17000g for 2 minutes at room temperature. The supernatant was discarded by pipetting and the pellet was dissolved with 1ml of 96-100 % ethanol, vortexed and centrifuged again for 2 minutes at room temperature. The supernatant was discarded; the tubes were air dried until all the residual ethanol was evaporated. The pellet was resuspended in 180µl buffer ATL followed by the addition of 20ul proteinase K, vortexed and incubated at 56°C for 1 hour. The samples were then incubated at 90°C for an additional 1 hour to make sure that the samples were completely lysed. The lysate was transferred to a QIAamp MinElute column (in a 2 ml collection tube), centrifuged at 6000g for 1 minute and the flow through was discarded. 500 µl Buffer AW1 was added to each QIAamp MinElute column, centrifuged at 6000g for 1 minute followed by an addition of 500 buffer AW2, and centrifuged again at 6000g for 1 minute. Each QIAamp MinElute column was centrifuged at 17000g for 3 minutes to remove any residual buffer from the column and transferred to clean 1.5 ml microcentrifuge tube. 20–100 µl of elution buffer (depending on the size of the pellet obtained) was added on each QIAamp MinElute column, incubated at room temperature for 5 minutes and

centrifuged at 17000g for 1 minute. The DNA was quantified using *Nanodrop 2000* spectrophotometer (Thermo Scientific, USA).

3.2.2 Extraction of tumour free circulating DNA from patients plasma

Tumour-free circulating DNA from the patients' serum was extracted using *ZR serum DNA kit* (Zymo research, USA). 2 ml plasma from each patient was transferred to a conical shaped 50 ml universal tube. 8 ml of genomic lysis buffer and 10µl of zymoBeads were added to and placed in a shaker for two hours at room temperature in order to mix the sample completely. The samples were centrifuged for 1-2 minutes; the pellet was resuspended in 500µl of DNA wash buffer, and transferred to a sterile 1.5 ml microcentrifuge tube. The samples were centrifuged at room temperature for 1 minute at 6000g, and the pellet was resuspended again in another 500µl of DNA wash buffer. The samples were centrifuged again at room temperature for 1 minute at 60000g to remove any residual buffer, and the pellet was air dried for 15 minutes. 30 µl of DNA elution buffer was added to the pellet (containing ZymoBeads), resuspended by repeated pipetting and was centrifuged at 10000g for 1 minute. The supernatant containing purified DNA was collected and was quantified using *Nanodrop2000* spectrophotometer (Thermo Scientific, USA).

3.2.3 Total RNA extraction

Total RNA was extracted using the *EZ-RNA extraction kit* (Biological Industries, Israel) from the fresh frozen metastatic brain tumours and breast cancer cell lines. Approximately 50 µg of the fresh frozen metastatic brain tumours were ground using

and pestle and mortar in dry ice, which were homogenized using 500 µl of denaturing solution to lyse the cells. The samples were vortexed and left at room temperature to make sure that the cells were lysed completely. 500 µl of extraction solution was added to each sample, vortexed and incubated at room temperature for 10 minutes followed by centrifugation at 17000g for 15 minutes at 4°C. This centrifugation step produced three phases containing RNA at the upper aqueous phase, DNA and protein in intermediate and lower phase respectively. The upper phase containing total RNA was carefully transferred to sterile RNAse free tubes, mixed with 500 µl of isopropanol, stored over night at -20°C and centrifuged at 17000g for 10 minutes at 4°C. The supernatant was removed; The RNA pellet was washed with 1 ml 75% alcohol (v/v), and centrifuged at 10000g for 5 minutes at 4°C. RNA pellet was air dried and dissolved with an appropriate amount (30-80 µl) of RNAse free water based on the size of the pellet.

To extract total RNA from the cell lines, respective cell lines were trypsinized using 1x 300µl trypsin to detach them from the flask when they were 80% confluent, and mixed with 1ml DMEM immediately to quench the trypsin. The cell suspension was centrifuged at 1500g for 5 minutes. The cell pellet was resuspended in PBS and re-centrifuged. RNA was extracted from the resulting pellet using the protocol for total RNA extraction.

Total RNA concentration was measured using a *Nanodrop2000* Spectrophotometer (Thermo Scientific, USA).

3.2.4 Quantification of nucleic acid preparations

The concentration of genomic DNA and total RNA was determined using Nano Drop 2000 Spectrophotometer (Thermo Scientific, USA). 2µl of water or the solution used to elute nucleic acids during preparation was used as a reference for zero absorbance (blank). 2µl of nucleic acid preparation was then used to quantify each sample. Nano Drop output readings consisted of a 260/280 and 260/230 ratio and a concentration in ng/µl. Purity of the DNA and RNA preparations was assessed using the 260/280 and 260/230 ratios. Both the DNA and RNA were considered to be pure if the ratio was in a range of ~1.8 to ~2.0. The ng/µl ratio was calculated by the spectrophotometer based on the modified Beer-Lambert equation *i.e.* c = (A*e)/b, where c denotes for nucleic acid concentration in ng/µl, A denotes for absorbance at 260nm, and e is the wavelength dependent extinction coefficient in ng-cm/µl for either double stranded DNA (50ng-cm/µl) or single stranded RNA (40ng-cm/µl).

3.2.4 Protein Extraction and quantification

Protein extraction was carried out from cell pellets derived from cancer cell lines. The cell pellets were lysed in RIPA buffer (25mM HCl, 0.1%SDS, 1% TritonX 100, 0.15M NaCl, pH 7.4) containing phosphatase and protease inhibitor (Roche, Germany). The lysate was transferred into a 1.5 ml microfuge tube, sonicated for 1 minute to further disrupt the cell wall and centrifuged at 4°C for 15 minutes. The supernatant containing protein was transferred to a new 1.5 ml microfuge tube.

Protein concentration was quantified using a modified Lowry method *i.e.* Bio-Rad detergent-compatible protein assay (Bio-Rad, United Kingdom). 5µl of each protein sample (in duplicates) was loaded to a 96-well plate. 20 µl of reagent S for every ml of

reagent A (an alkaline tartrate solution) was mixed in a separate tube, and 25 µl of that mix was added into each protein sample. 200 µl of reagent B (A dilute Folin reagent) was added to each protein sample and the plate was incubated at room temperature for 10 minutes to allow colorimetric reaction. The plate was read on a spectrophotometer (Thermo Scientific, USA) at 650nm.

3.3 DNA Methylation Analyses

3.3.1. DNA methyltransferase modification of DNA to create fully methylated positive controls

Fully methylated positive control DNA was generated by incubating genomic DNA with DNA methyltransferase in the presence of S-adenosyl methionine (SAM) (New England Biolab, USA). SAM acts as a donor (source) for a methyl group whereas DNA methyltransferase (DNMTs) catalyzes the transfer of methyl group to cytosine residues. The 50µl of reaction mixture contained 5µg of genomic DNA, 5 µl NEB buffer, 2µl of 32mM SAM, 2µl (4 units) of DNMT and sterile distilled water. The reaction mixture was incubated for 2 hours at 37°C resulting in 100ng/µl of fully methylated DNA.

3.3.2 Bisulphite Conversion of DNA

Bisulphite conversion of genomic DNA from tumour samples was carried out using *EZ DNA methylation kit* (Zymo Research Corp. USA) whereas, Bisulphite conversion of tumour free DNA from plasma was carried out using EpiTect bisulphite kit (Qiagen, USA)

3.3.2.1 Bisulphite conversion of genomic DNA

500ng of genomic DNA and 1000ng (1µg) of of fully methylated positive controls were bisulphite converted using the *EZ DNA methylation kit* (Zymo Research Corp. USA). 5µl of M-dilution buffer and water was added to each sample to make a total volume of 50µl. The samples were incubated at 37°C for 15 minutes. 100µl of CT conversion reagent was added to each sample, mixed by pipetting and were incubated at 50°C for 12-16 hours. The samples were kept on ice for 10 minutes after the incubation was over. 400µl of M-binding buffer applied to each zymo spin column with collection tube, and the samples were added and mixed by inverting the column 4-6 times. The columns were centrifuged, washed using 100µl of M-wash buffer and centrifuged again. 200µl of M-disulfonation buffer was added to each column, incubated at room temperature for 15-20 minutes and centrifuged to ensure the removal of excess bisulphite salts by disulfonation. Each column was washed twice, using 200µl of M-wash buffer and centrifuged at full speed for 1 minute. The bisulphite converted DNA was eluted using 30 µl of elution buffer and was stored at -20°C.

3.3.2.2 Bisulphite conversion of tumour free circulating DNA from patients' plasma

Epitect Bisulfite kit (Qiagen, USA) was used for bisulphite conversion of tumour-free circulating DNA from patients' plasma. 200ng of serum or plasma DNA were added to 200µl sterile PCR tube, the final volume was made up to 20µl with RNase free water.

85µl bisulphite conversion reagent (bisulphite mix) and 85µl DNA protect buffer was added to the samples making up total volume of 140 µl.

The samples were incubated in a thermal cycler for three rounds of denaturation and inubation at 95°C and 60°C respectively. Each denaturation steps consisted of 5 minutes whereas incubation time varied in each cycle. First round of denaturation and incubation consisted of 5 minutes and 25 minutes, 2nd round consisted of 5 minutes and 85 minutes and where as third cycle consisted of 5 minutes and 175 minutes (2 hour 55 minutes) respectively. The bisulphite converted DNA was transferred to a sterile 1.5 ml microcentrifuge tube and 560µl of freshly prepared buffer BL containing 10 µg/ml carrier RNA was added to each tube in order to ensure binding of DNA with the column. The samples were transferred to EpiTect spin columns with collection tubes and were centrifuged for 1 minute at room temperature. The columns were washed with 500 µl of buffer BW and the sample were desulfonated by adding 500µl of buffer BD to each column, incubated at room temperature for 15 minutes, and were centrifuged for 1 minute at room temperature. The columns were washed twice by the addition of 500 µl of buffer BW and centrifugation for 1 minute at room temperature. The columns were transferred to a new sterile 1.5 ml microcentrifuge tube, 20µl buffer EB was added to each column and centrifuged for 1 minute in order to elute DNA. Another 20 µl buffer EB was also added to column and centrifuged again to increase the yield of DNA eluted.

3.3.3 Combined Bisulphite and Restriction Analysis

3.3.3.1 CoBRA primer design

CoBRA PCR was carried out using semi nested and nested primers. The primers were designed to amplify up to 550bp of a promoter region CpG Island. CpG islands associated with candidate genes were identified on the human genome browser (http://genome.ucsc.edu/cgi-bin/hgGateway). Primers were designed using criteria to maximize the specificity and effectiveness of the primers designed; (i) The 5'end of the primers preferably started with a G nucleotide and the 3' end of the primers preferably ended with ACC, or CAC in unmodified DNA resulting in ATT or TAT in bisulphite modified DNA in order to provide a strong initial binding point and a high specificity to bisulphite modified DNA. ii) Primers didnot preferably contain strings of As, Ts, or Cs in them to prevent formation of hairpin structures. It was important to ensure that CoBRA primers were unbiased towards methylation status. The incorporation of CG dinucleotide within the primer was avoided wherever possible to ensure that the primers designed were unbiased on methylation status of the region amplified. However if this was unavoidable, C nucleotides were designated as Y (either C if methylated or T if not methylated, to make a 50:50 ratio of C or T nucleotide incorporation) in forward primers. Similarly, C nucleotide was designated as R (either G, if methylated or A if not methylated, to make a 50:50 ratio of G or A nucleotide incorporation) in reverse primers. The incorporation of BstUI restriction sites (CGCG) or TaqI restriction sites (TCGA) within primers were avoided. Annealing temperatures of primers were calculated using the equation: [(% C or G x 0.41) + 64.9 - (600/N)] where % C or G is calculated from the total number of C or G in the primer (Y and R is counted as 0.5) and N is the total

number of nucleotides with in the primer (length of the primer). All primers were manually designed.

The CoBRA primers designed are listed in appendix A3.1.

3.3.3.2 CoBRA PCR

Semi nested or nested primers were used in all CoBRA PCR reactions, therefore, two rounds of PCR were carried out per PCR. Both the nested PCR (first round) and semi-nested (second round) reactions contained 2.5 µl 10X PCR buffer containing magnesium chloride (pH 8.3), 2.5mM dNTP mix (Bioline, UK), 20µM of forward and reverse primer and 6.25 units or 7.5 units (for primary or secondary reactions respectively) of FastStart Taq DNA polymerase (Roche, Germany). Reactions were made up to 25µl or 50µl for primary and secondary reactions respectively with distilled water. 2µl of bisulphite converted DNA was used as starting material in the first nested/seminested PCR and 2µl of this PCR product was used as starting material for the second round of nested/semi nested PCR. The reagents and concentration of nested and semi nested PCR is given in table 3.3.

Reagent	Initial Concentration	Volume used First round	Volume used Second round
PCR buffer with $MgCl_2$	10X	2.5µl	5µl
dNTPs	2.5 mM	2.5µl	5µl
Forward Primer	20µM	1.25µl	1.5µl
Reverse Primer	20µM	1.25µl	1.5µl
Taq polymerase	5 U/µl	0.25µl	0.25µl
DNA	Bisulfite converted DNA	2µl	2µl from the first round
Double-distilled water	-	15.3µl	34.75µl
Total	**-**	**25µl**	**50µl**

Table 3.3: Reagents used to set up nested and semi nested PCR reaction. Master mix was prepared with the required reagents. 2 µl of bisulphite converted DNA (for the first round) was transferred into a sterile 200 µl PCR tube and 23 µl of master mix was added on the DNA. For the semi-nested PCR, 2 µl of nested PCR product was transferred to a new sterile 200 µl PCR tube followed by 48 µl master mix

Both nested and semi-nested PCR were carried out using touchdown PCR, see tables 3.4 for program details.

Temperature	Duration	Step	Description/cycle
94°C	3 minutes	Initial denaturation	
94°C	45 seconds	Denaturation	2 cycles
Tm +4°C	45 seconds	Annealing	
72°C	45 seconds	Extension	
94°C	45 seconds	Denaturation	2 cycles
Tm +2°C	45 seconds	Annealing	
72°C	45 seconds	Extension	
94°C	45 seconds	Denaturation	2cycles
Tm	45 seconds	Annealing	
72°C	45 seconds	Extension	
94°C	45 seconds	Denaturation	**32 cycles**
Tm -2°C (TD)	45 seconds	Annealing	
72°C	45 seconds	Extension	
72°C	**5 minutes**	**Final extension**	

Table 3.4: Touchdown PCR program. Touchdown PCR program starts at a higher temperature than the annealing temperature (Tm) of the primers. The program was designed to run first two cycles at Tm+4°C, the next two cycles at Tm+2°C, the next two cycles at Tm. Finally, the reaction mixture was run for 32 cycles at touchdown temperature (Tm-2°C). .

3.3.4. Methylation Specific PCR (MSP)

3.3.4.1. MSP primer design

MSP primers were designed for those genes which methylation status suggested clinical significance and deserved further investigation. MSP product sizes ranged from 120-250bp in length and were located within the region analysed by CoBRA (to avoid biases in methylation analysis of the same region). For MSP, two sets of forward and

reverse primers were designed *i.e.* Unmethylated Specific PCR (USP) and MSP primers. USP primers were specific for unmethylated sequences and MSP primers were specific for methylated sequences. They are designed with the following criteria and guidelines for designing MSP primers; i) In contrast to CoBRA primers, MSP primer sequence contained at least 2-3 CG dinucleotide to provide specificity and to avoid biases in methylation status analysis., ii) Primers were 25-35 base pairs in length iv), As there are less C or Gs in USP primers than MSP primers, they are often longer to accommodate additional Gs in their 5'end to ensure their annealing temperature is in a same range as in MSP. In addition, primer-designing criteria from CoBRA primers were considered (section 3.3.3.1). All primers were manually designed.

The MSP primers designed are listed in appendix A3.2.

3.3.4.2 MSP in brain metastases and patients serum

MSP was used to analyse free circulating DNA isolated from patients' serum and brain metastases. MSP in brain metastases was carried out to ensure that the MSP primers designed for each gene are able to amplify the region of interested before carrying out MSP/QMSP in patients' serum. Each 25µl MSP reaction contained 2.5µl of 10X PCR buffer containing magnesium chloride (pH 8.3), 2.5mM dNTPs mix (Bioline, UK), 20µM forward and reverse primer each and 6.25 units of FastStart Taq DNA polymerase (Roche, Germany). (Table 3.3). A touch down PCR program was used to amplify the sample using either MSP or USP primers (table 3.4).

3.3.5 Quantitative Methylation Analysis by Bisulphite Sequencing

3.3.5.1 Purification of PCR Products

PCR products selected for cloning were first purified from agarose gels using a Gel Extraction Kit (Qiagen, USA). PCR products were run on 1% agarose gel electrophoresis. The products were cut out of the gel using a scalpel and transferred to 1.5ml microcentrifuge tubes. Gel slices were weighed and 3 times mass (mg) in volume (µl) of buffer QG was added to samples and was incubated at 50℃ for 10 minutes with intermittent vortexing in every 2-3 minutes to dissolve the agarose gel. 1 x mass (mg) in volume (µl) of 100% isopropanol was added to the solution, mixed and was transferred to transferred to a QIAquick spin column and centrifuged at room temperature for 1 minute at 17000g. The flow-through was discarded, 500µl buffer QG was added and was centrifuged again. Flow-through was again discarded, 500µl buffer PE was added to the column and the samples were centrifuged again for 1 minute at 17000g at room temperature. The samples were centrifuged again to remove residual ethanol. The columns were placed in a fresh 1.5ml microcentrifuge tube, 30µl distilled water was added directly to the column membrane and was incubated at room temperature for 1 minute. The tubes were centrifuged at room temperature to elute the purified DNA from the column. The purified PCR product was either used immediately or stored at -20°C.

3.3.5.2 Ligation of PCR products into plasmid vector

The PCR products were ligated in the pGEM-T Easy vector system (Promega, USA) in order to introduce PCR products into bacterial cells for cloning. The ligation reaction was set up with 0.5µl pGEM vector, 1µl T4 DNA ligase and 5µl 2X ligation buffer. This was then added to 3.5µl purified PCR product. Samples were left to incubate at 4℃ overnight for maximum ligation efficiency.

3.5.5.3 Preparation of ampicillin containing (Amp+) agar plates and LB Broth

25g of LB (Luria-Bertani) medium (5N NaCl, tryptone and yeast extract, pH 7) (Sigma, UK) and 15g of agar were dissolved in 1 litre distilled water and were autoclaved. Ampicillin was mixed to give a final concentration of 100µg/ml. Agar was then immediately poured into the sterile petri dishes, allowed to cool and set, and were transferred to 4°C for storage. For LB broth, 25g/L of LB was dissolved in distilled water without agar and was autoclaved, it was stored at 4°C. a small aliquot was transferred to a sterile tube and ampicillin was added to it to give a final concentration of 100µg/ml.

3.5.5.4 Transformation of PCR product ligated into the pGEM-T easy vector into the bacterial cells

Ligated PCR products were mixed with 50µl DH5α competent cells (Life Technologies, UK) by gentle shaking. The cells were heat shocked for 1 minute at 42℃ and were then incubated on ice for 5 minutes. 1 ml of SOC media (2% w/v, trypton, 0.5%w/v yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 20 mM glucose) (Invitrogen, USA) was added to the cells, which, were then incubated for one hour at 37℃. The cells were centrifuged at 2000g for 3 minutes, and two-third of supernatant was gently drained off from the tube. The pellet was resuspended in the remaining supernatant. These cells were plated on Amp+ LB agar plate and were incubated overnight (or up to 16 hours at 37℃).

3.5.5.5 Single colony PCR

To ensure the greatest chance of picking colonies containing the insert, only white colonies were picked, transferred to 20µl distilled water and heated to 95℃ for 5

minutes to rupture all the cellular membranes. Once cooled, 7µl was used as starting material in a PCR reaction (see table 3.3 for protocol), which contained 1x 5X GC rich solution, 25µl with distilled water. The Single-colony PCR reactions were carried out using a touchdown PCR program (table 3.4)

3.3.5.6 PCR product clean-up for sequencing

Single colony PCR products were cleaned-up using a combination of FastAP alkaline phosphatase (AP) (Fermentas, USA) to remove excess dNTPs, and ExonucleaseI (ExoI) (New England Biolabs, USA) to remove excess single stranded primer. A reaction mixture contained 10µl PCR product, 1U FastAP, 0.01U ExoI and 10x FastAP buffer (100mM Tris-HCL (pH8), 50mM magnesium chloride, 1M potassium chloride, 0.02% Triton X-100 and 1mg/ml bovine serum albumin (BSA)). Samples were incubated for 30 minutes at 37℃ followed by an enzyme inactivation step for 20 minutes at 80℃.

3.5.5.7 Sequencing reaction

Sequencing reactions were carried out using Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, UK). 5µl cleaned up PCR product was added to a 20 pmol forward or reverse primer, 0.5µl Big Dye, 2µl 5X buffer and made up to a total volume of 10µl with distilled water. The sequencing reaction contained initial denaturation at 94℃ for four minutes followed by 35 cycles, each of which contained three temperatures (94℃ for 25 seconds, 50℃ for 25 seconds and 60℃ for four minutes). Samples were wrapped in foil following the sequencing reaction and stored at -20℃ until precipitated and prepared to be read on the sequencer.

3.5.5.8 Ethanol precipitation following sequencing reaction

After the sequencing reaction, 3.5µl precipitation buffer (1.5M sodium acetate, 1.5M EDTA) and 100µl of 95% (v/v) ethanol and were centrifuged at 2,254g for 30 minutes at 4℃. The supernatant was carefully removed and the samples were then briefly pulsed at 23g upside down and were washed with 200µl 70% (v/v) ethanol and centrifuged again at 2,254g for 30 minutes at 4℃ for. The supernatant was removed and samples were pulsed again at 23g upside down, washed with 200µl 70% (v/v) ethanol and were centrifuged again at 2,254g for 30 minutes at 4℃. The supernatant was removed and samples pulsed at 23g upside down and were allowed to air dry.

For sequencing, pellets were resuspended in 10µl Hi-Di Formamide (Applied Biosystems, UK) and denatured at 95℃ for 5 minutes. Samples were run on a 3730 DNA Analyzer (Applied Biosystems, UK) and output files analysed using sequencing analysis 5.2 (Applied Biosystems, UK).

3.4 Expression analyses of selected genes

3.4.1 Synthesis of complementary DNA (cDNA) from total RNA

Complementary DNA (cDNA) synthesis was carried out using high fidelity cDNA kit (Roche, Germany). 1µg of total RNA was mixed with 1µl of oligos dT primers (50pmol/µl) and PCR grade water to make 11.6µl volume. The samples were incubated at 50°C PCR grade water to make 11.6µl volume. During the incubation time, 8.6µl of master mix containing 0.5µl (20 units) of RNAse inhibitor, 2µl of dNTP mix (10mM), 1µl DDT (5mM), 1.1µl of reverse transcriptase (22 units) were prepared and added to each sample to make total volume of 20µl. The samples were incubated at 50°C for 30

minutes followed by 85°C for 5 minutes, samples were then immediately placed an ice to stop the reaction.

3.4.2 Reverse transcription polymerase chain reaction (RT-PCR) primer design

RT primers were designed from primary transcripts of the genes of interest. Exon sequences of candidate genes were retrieved from Ensembl Genome Browser (http://www.ensembl.org/index.html). RT primers were designed in such a way that the product encompasses more than one exon, thus ensuring the amplification product was not derived from any contaminating gDNA. Primers were designed manually within certain guidelines; 1) primers were 18-25 base pairs long (mostly around 20 base pairs), 2) primers start with G or C to ensure binding specificity of the primers with template cDNA, 3) there is even distribution of C and G nucleotide as C and G nucleotide determines annealing temperature of the primers which affects the annealing of the primers with the template cDNA, 4) long stretches of any nucleotide is avoided to prevent formation of hairpin or secondary structure, 5) the product size is around 300-500 base pairs. Annealing temperatures of primers were calculated using the equation: [(% CG x 0.41) + 64.9 - (600/N)] where % CG is calculated from the total number of C and G in the primer and N is the total number of nucleotides with in the primer (length of the primer). All primers were manually designed.

The RT primers designed are listed in appendix A3.3.

3.4.3 Reverse transcription (RT) PCR

1µl of prepared cDNA was used to carry out RT PCR to investigate the expression of genes in breast cancer cell lines and in metastatic brain tumours. 25µl volume of RT-PCR reaction contained 2.5µl 10X PCR buffer containing $MgCl_2$ (pH 8.3), 2.5mM

dNTPs mix, 1.25 μM forward primer, 1.25 μM reverse primer, 0.5U Fast Start Taq DNA polymerase (Roche, Germany), and 16.3 μl distilled water. RT-PCR reactions were carried out using a touchdown PCR (3.4). The β-actin gene was used as a positive housekeeping control with the same conditions except that the number of cycles was reduced (20 cycles) to compare the expression level of the samples being used (28 cycles).

3.4.4 Quantification of gene expression

Quantification of gene expression was carried out using a densitometry program from the genetool software to investigate if methylation correlates with expression for three genes (*BNC1*, *CCDC8* and *GALNT9*). Each of the RT PCR products for respective gene run on agarose gel electrophoresis was quantified using densitometry on genetool software. The PCR product for each sample was also quantified for β-actin gene. The expression level of each sample for respective genes was calculated by comparison with expression level of β-actin gene.

3.4.5 Western blotting

Protein extraction and western blotting of extracts from breast cancer cell lines transfected with control oligos/siRNA against candidate genes or transfection ready DNA (cDNA clone) for *BNC1*, *CCDC8* and *GALNT9* were performed. The western blot was performed as follows.

Sample preparation and loading in a gel:

70μg of each protein extract (section 3.2.4, for protein extraction and quantification) was transferred to 0.5ml microcentrifuge tube and 6μl 4x loading buffer, (106 mM Tris

HCl, 141 mM Tris Base, 2% SDS, 10% Glycerol, 0.51 mM EDTA, 0.22 mM SERVA Blue G250, 0.175 mM Phenol Red at pH 8.5), 1µl DDT and water was added to make a final volume of 25µl. The samples were placed at 96°C for 10 minutes in order to denature protein.

Samples were loaded on 6-10% polyacrylamide gels (PAG) based on the protein size. Each gel consisted of a resolving and a stacking gel. 6% resolving polyacrylamide gels were used for the proteins above 100 KD and 10% polyacrylamide gels were used for the proteins below 100 KD. A polyacrylamide resolving gel was poured between mini western mini apparatus (Biorad, UK) to a depth of ~5.0 cm, topped with water to get a uniform surface and was left to polymerise for 10-15 minutes. Stacking gel was poured to the top of the plates, a 1mm comb with 10 lanes was placed immediately and the gel was left to polymerise for another 10-15 minutes. The samples prepared were loaded and gel was run in 1x SDS PAGE running buffer (Gene flow, UK) at 200V for 60-75 minutes or until the smallest marker (11KD) on the protein ladder (geneflow, UK) reached to the bottom of the gel in a vertical electrophoresis unit.

The gel was removed from a glass plate carefully and was placed in 1x transfer buffer (Geneflow, UK) containing 20% methanol (v/v). Polyvinyl difloride (PVDF) transfer membrane (GE healthcare, United Kingdom) was first soaked in methanol followed by 1x transfer buffer. The gel was placed in a membrane sandwiched between 3 sheets of 3MM blotting papers on each side. Transfer was carried out in a blotting Unit for 1.5 hours at 20V and 200mA.

Probing with antibody

The membrane was then blocked in blocking buffer (1x TBS, 1% Tween and 5% milk) for 30 minutes followed by probing with rabbit primary antibody against the protein being analysed at 1-2.5μg/mL (Table 3.5) at 4°C overnight. The membrane was washed twice with 1xTBS-T *i.e.* containing 1% tween X-100 (Sigma, UK) for 10 minutes. Signals were detected by probing the membrane using horseradish peroxidase–conjugated anti-rabbit antibody (GE healthcare, United Kingdom) at 1μl/5mL in 5% milk at room temperature for 2 hours. The membrane was washed again for 10 minutes twice in 1x TBST. Equal volume of ECL reagent A (luminol) and B (enhancer) was mixed from an ECL Chemoluminescence kit (Biological Industries, Israel) and applied to a membrane. The membrane was incubated at room temperature for 5 minutes; excess reagent was drained off and was covered by cling film in a cassette. The image was exposed to a Carestream Kodak BioMax MR film (Sigma, UK) in a dark room for a suitable time period (protein dependent), generally 5-15 minutes and immediately dipped in developer solution for up to 1 minute. The image was then dipped in a fixative solution to fix the image. The membranes were again washed in 1xTBST for 10 minutes and stained with 1μl/1ml (v/v) India ink (Winsor and Newton, United Kingdom) for comparison of loading.

Protein	primay antibody conc.	secondary antibody conc.	exposure time	Molecular weight (KD)
BNC1	2.5μg/ml	2μL/ml	20 minutes	105
CCDC8	1μg/ml	1μL/ml	10 minutes	64
GALNT9	1μg/ml	1μL/ml	15 minutes	68

Table 3.5: Concentration of primary antibody and secondary antibody, exposure time and molecular weight of proteins BNC1, CCDC8 and GALNT9 analysed by

western blotting. Generally, higher concentration of antibodies and exposure time were needed for BNC1 compared to CCDC8 and GALNT9.

3.5 Tissue Culture

A class-two laminar airflow hood was used to carry out tissue culture using sterile technique at all times. Laminar airflow hoods were cleaned using 1x trigene followed by 70% ethanol (v/v). All other equipment was wiped down with 70% ethanol (v/v) before use in the hood. Five breast cancer cell lines (see section 3.1.5) were grown in this study.

3.5.1 Preparation of growth media

Breast cancer cell lines used for this study were routinely maintained in Dulbecco's Modified Eagles Media (DMEM) (Sigma, UK) containing L-Glutamine supplemented with 10% (v/v) fetal bovine serum (FBS) (Thermo scientific, USA), penicillin and streptomycin (100u/ml (Lonza, Switzerland).

3.5.2 Revival of cell lines

Cell lines frozen stored in liquid nitrogen were removed and defrosted carefully and were transferred to 12-15 ml pre-warmed (37℃) DMEM in T75 (75cm^2) tissue culture flask supplemented with 10% FCS, 5% CO_2, and 500 U/ml penicillin-streptomycin. Freshly revived cells were maintained at 37℃, 5% carbon dioxide.

3.5.3 Maintenance and passaging of cell lines

Medium from the cell lines were replaced with fresh pre warmed medium every 3-4 days. The medium from the flask was first removed, cells washed with 10 ml buffered

saline (PBS) and fresh medium was added to the flask. Cell lines were passaged into a new flask with a ratio of 1:3, 1:5 or 1:10 when they reached approximately 70%-90% confluency depending on the growth rate of the cell lines or the requirement to carry out further functional work. After the removal of media, cell lines were washed with 10 ml PBS, and 2 ml pre-warmed 1x Trypsin-EDTA (Lonza, Switzerland) was added to the flask drop by drop ensuring spreading and reaching of trypsin uniformly in order to detach the cells from the flask. After 1-2 minute incubation or at 37℃ when the cells were completely dislodged from the flask surface, cells were resuspended in 10 ml media and split with the desired ratio to ensure the right volume of cells were maintained in the flask. The unwanted cells were discarded and the cells in the flask were topped up with media to a final volume of 15 ml.

3.5.4 Counting of cells using a Haemocytometer

Cells were fully trypsinised as described in section 3.5.3 and resuspended in 8 ml fresh media. A coverslip was placed firmly onto the Haemocytometer and 10µl of cell suspension was applied to both sides of the heaemocytomer underneath the coverslip ensuring the cells suspension was uniformly placed in the Haemocytometer chamber. The cells in each side were counted under a microscope and an average cell count was taken.

3.5.5 Treatment of cell lines with 5-aza-2'deoxycitidine

5-AZA-2'deoxycitidne (5-AZA-dC), a global demethylation agent, was used to demethylate cell line genomic DNA to enable comparison of methylation status of selected genes. 5'Aza-dC incorporates methyl groups to genomic DNA during its replication in dividing cells. Therefore, routinely maintained five breast cancer cell lines

were plated to make 30-50% confluent based on their doubling time to ensure that both control and 5'AZA-dC-treated cells lines were approximately 75 % confluent at the time of RNA extraction. 5-AZA-dC was freshly prepared in double distilled water and filter sterilized. 24 hours after seeding, cells were treated with 5μM 5–AZA-dC supplemented medium. Cells were treated with fresh 5μM 5-AZA-dC three times a week on alternate days. After 7 days, cells were washed with PBS and cell pellets prepared as describe in section 3.5.6

3.5.6 Preparation of cell pellets

To prepare the cell pellet, medium from the cells was aspirated, washed with PBS and trypsinised. After the cells were fully dislodged, 8 ml of medium was added to suspend the cells. Cells were resuspended in culture media, transferred to a 25ml universal tube and were centrifuged at 1200g at room temperature for 5 minutes. Supernatant was aspirated carefully; cell pellets were resuspended with 3ml PBS (equal volume of media) and centrifuged at 1200g for 5 minutes. The supernatant was removed and the pellets were immediately placed on ice for DNA/RNA/Protein extraction. For long-term storage, pellets were snap frozen in liquid nitrogen before transferring to -80℃ for long-term storage.

3.5.7 Freezing-down cell lines

After the cells reached 80-90% confluency, they were first washed with PBS, trypsinised, and resuspended in 8ml of media described in sections 3.5.3 and 3.5.6. The cells were then transferred to a 25ml universal tube, centrifuged at 1200g at room temperature for 5 minutes, washed with PBS and centrifuged again. The cell pellet was resuspended in 1 ml freezing buffer (10% dimethyl sulphoxide, v/v (DMSO) in FBS) to

act as a cryopreservant during storage. The cell suspension was transferred to a cryo-vial, wrapped in tissue paper and immediately transferred to -80℃, such that under these conditions the temperature goes down at approximately a -1℃ per minute. Vials were transferred to liquid nitrogen for long-term storage.

3.6 Functional Analysis of Genes/Proteins

3.6.1 RNA Interference (RNAi) assays and over-expression of genes in breast cancer cell lines

RNAi was carried out to transiently knockdown genes using siRNA nucleotide oligos (Ambion, USA) in breast cancer cell lines. The breast cancer cell lines T47D, MCF7 or MDA-MB231 were washed, trypsinised and counted as described in section 3.5.4. Cells were seeded into six-well plates containing DMEM with 10 % FBS at a density of 0.5x 10^6. After overnight incubation, candidate genes were knocked down by transfection of RNAi 'silencer select' oligos against respective genes. Control cells were transfected with 'silencer select' control Oligo no. 1 (Ambion, USA). For each well, 250 µl of optimem media and 4µl of Lipofectamine 2000 reagent (Life Technologies, UK) were mixed in a 10ml tube and, 250µl of optimem medium and 20 pmol of siRNA oligos were mixed in another 10ml tube. After five minutes incubation at RT, both the tubes were mixed together. This mixture was incubated for 15-20 minutes at room temperature. During the incubation time, the cells were washed with PBS and media replaced with 1 ml of Optimem medium. The reaction mix was added to each well drop-by-drop and shaking forward and backward. The plated cells were returned back the incubator. After 4-6 hours, 1ml of antibiotic free DMEM medium was added to each well and returned to the incubator (37°C). 48 hours after knockdown, cells were

collected and isolated as described in section 3.5.6. The transfection/knockdown was carried out in triplicates. Total RNA and protein was extracted to validate knockdown of genes and proteins using RT-PCR and western blotting.

3.6.2 Migration assay

The selected genes/proteins from the breast tumour cell lines were knocked down using siRNA oligos as described in section 3.6.2. 24 hours after the transfection, DMEM with 10 % FBS was replaced with fresh DMEM without FBS and incubated at 37°C for another 24 hours. The confluent monolayer of cells in each well was "scratched" with the tip of a 200 μl pipette. The cells were washed with PBS and replaced with the fresh serum free media. Images were taken of each well/"scratch" this was considered the beginning of the experiment and referred to as 0 hour. The extent of migration of cells was observed after 24 hours and 48 hours of initial scratching. The regions captured were marked to ensure that the same region is observed after 24 hours and 48 hours under the microscope to avoid biasness. In addition, similar transfection experiments were set up in parallel and cells were collected after 48 hours to perform RT-PCR and Western blot analysis to determine the knockdown efficiency in these cells. The distance migrated by cells towards the gap was examined using Image J software (Schneider *et al.*, 2012). Assays were performed in triplicate for each knockdown to minimize experimental biases. Statistical analyses were carried out using Microsoft Excel and p values were calculated using t tests from graphpad prism as described in section 3.8.1.

3.6.3 Invasion assay

Invasion assays were carried out to determine the invasive potential of the cancer cell lines following knockdown and silencing of candidate genes. This was done by trans-well assay using 24 well plates containing inserts with 8μm pores (Corning, USA). Knockdown was carried out as described in section 3.6.2. 24 hours after the knockdown, DMEM with 10 % FBS was replaced with fresh DMEM without FBS and incubated at 37°C for another 24 hours. 200μl of matrigel matrix (Becton Dickinson Labware, USA) diluted 10-fold with DMEM without FBS was applied to 24-well 9-mm inserts containing polyethylene terephthalate (PET) membranes with 8μm pores (Corning, USA). The inserts were placed in sterile 24 well plates and incubated at 37°C for 30 min to settle. During the incubation period, the breast cancer cell lines knocked down with respective oligos were harvested and resuspended in serum free medium and counted using a haemocytometer as described in section 3.5.4. The cell density was maintained at $1x10^5$/ml with addition of appropriate volume of serum free media in each tube to minimize experimental biasness. 300μl of the cells (150,000 cells) were applied to the matrigel invasion chamber. 500 μl of DMEM containing 10% FBS was placed in the lower chamber (in 24-well plate) as a chemo-attractant. The plates were incubated at 37°C for 48 hours with 5% CO_2. After 48 hours, medium from the lower and upper chamber was carefully removed without disturbing the cells. Cells from the upper surface were removed by wiping with a cotton bud and the inserts were transferred to a new 24 well plate containing 400μl crystal violet and were incubated for 10 minutes to stain invasive cells on the lower layer. The inserts were washed with water, air-dried at room temperature and three fields from each insert were counted at 200X magnification. Extraction solution (Cell biolabs, USA) was applied to the plates and left for 10 minutes. 100μl of cell solution was transferred to fresh 96-well plate to determine the

optical density of each well at 540 nm. Assays were performed in triplicates for each knockdown. Statistical analyses were carried out as described in section 3.8.1.

3.7. Illumina Infinium HumanMethylation450k Data Analyses

3.7.1 Analyses of the 450K Methylation data from The Cancer Genome Atlas

In order to identify novel candidate metastatic suppressor genes, Illumina HumanMethylation450K BeadChip array data from The Cancer Genome Atlas (TCGA) was analysed. To ensure that the genuine promoter-associated CpG islands only those probes that are located in the 5' region of the gene or up to 1500 base pairs from the transcription start site (TSS, TSS200, TSS1500) and 5' UTR were selected. Methylation array data for 20 primary lung tumours and 20 primary breast tumours (with no evidence of metastasis) were downloaded from the TCGA (Table 3.6) and were compared for the methylation status of individual probes across the genome. Comparison was carried out between individual probes that are not methylated (β value ≤ 0.25) in 75% (15/20) in primary breast tumours and methylated (β value ≥ 0.60) in primary lung tumours, in at least 50% (10/20) of the samples. Common genes in these lists were identified. This analysis identified four candidates that were frequently methylated in lung tumours and infrequently methylated in breast tumours. The details of the probes and genes are explained in section 5.

BRCA		LUAD	
Tumour Code	Tumour barcode	Tumour Code	Tumour barcode
B1	TCGA-AC-A23H-01A-11D-A161-05	L1	TCGA-38-4631-01A-01D-1756-05
B2	TCGA-BH-A204-01A-11D-A161-05	L2	TCGA-38-4632-01A-01D-1756-05
B3	TCGA-BH-A208-01A-11D-A161-05	L3	TCGA-44-6144-01A-11D-1756-05
B4	TCGA-BH-A209-01A-11D-A161-05	L4	TCGA-44-6145-01A-11D-1756-05
B5	TCGA-E2-A1LI-01A-12D-A161-05	L5	TCGA-44-6146-01A-11D-1756-05
B6	TCGA-E2-A1LS-01A-12D-A161-05	L6	TCGA-44-6147-01A-11D-1756-05
B7	TCGA-E9-A1RB-01A-11D-A161-05	L7	TCGA-44-6148-01A-11D-1756-05
B8	TCGA-E9-A1RC-01A-11D-A161-05	L8	TCGA-49-4488-01A-01D-1756-05
B9	TCGA-E9-A1RD-01A-11D-A161-05	L9	TCGA-50-5930-01A-11D-1756-05
B10	TCGA-E9-A1RF-01A-11D-A161-05	L10	TCGA-50-5931-01A-11D-1756-05
B11	TCGA-A2-A3KC-01A-11D-A212-05	L11	TCGA-53-7813-01A-11D-2168-05
B12	TCGA-A2-A3KD-01A-12D-A212-05	L12	TCGA-55-7724-01A-11D-2168-05
B13	TCGA-A7-A3IZ-01A-11D-A212-05	L13	TCGA-55-7725-01A-11D-2168-05
B14	TCGA-A7-A3J0-01A-11D-A212-05	L14	TCGA-55-7726-01A-11D-2168-05
B15	TCGA-A7-A3J1-01A-11D-A212-05	L15	TCGA-55-7727-01A-11D-2168-05
B16	TCGA-AC-A23G-01A-11D-A212-05	L16	TCGA-55-7815-01A-11D-2168-05
B17	TCGA-AC-A3HN-01A-11D-A212-05	L17	TCGA-55-7816-01A-11D-2168-05
B18	TCGA-C8-A3M8-01A-11D-A212-05	L18	TCGA-55-7903-01A-11D-2168-05
B19	TCGA-E2-A3DX-01A-21D-A212-05	L19	TCGA-55-7907-01A-11D-2168-05
B20	TCGA-E9-A3HO-01A-11D-A212-05	L20	TCGA-55-7910-01A-11D-2168-05

Table 3.6: A list of primary breast tumours (BRCA) and primary lung tumours (LUAD) downloaded from The Cancer Genome Atlas (TCGA) to carry out bioinformatic analyses of genome wide 450K methylation data to identify novel candidate genes that may contribute to breast cancer metastases to brain (BBM). The details of the analyses are given in section 5.

3.7.2 Analyses of the 450K Methylation data for breast cancer brain metastases

Genomic DNA from 24 breast to brain metastases (BM1-BM30) was extracted (see section 3.2.1.1) and prepared at a final concentration of 50ng/μl. This DNA was used to interrogate Illumina Infinium 450K methylation Beadchip arrays. Bisulphite modification of these samples, array hybridisation and preliminary data analysis was carried out by Cambridge genome services, United Kingdom. Raw data was provided as

β-values that represented methylation levels of ~485K genomic probe-hybridising regions. β-values range from 0 (designating no methylation) to 1 (Designating complete methylation). β-values of each individual probes for breast to brain metastases samples were compared to normal breast tissues and primary breast tumours from TCGA to identify probes which are either hypomethylated or hypermethylated in breast metastases compared to normal breast tissues and primary breast tumours.

3.8 Statistical Methods:

3.8.1 Fisher's exact test, t test and p value calculation

Fisher's exact test was used to determine the significance of the frequency of methylated samples in a particular group such as the primary breast tumours or the metastatic brain tumours from the graphpad prism (http://graphpad.com). In Fisher's exact test, $p<0.05$ was considered statistically significant. Initially, Fisher's exact test was carried out to investigate the statistical significance of methylation status between primary breast tumours (n=20) and BBM (n=15), and those genes, which are statistically significant, is further analysed in another cohort of primary breast tumours (n=40) and BBM (n=30). Similarly, t test was used to calculate p value in migration and invasion assay.

3.8.2 Kaplan-Meier analysis

Kaplan-Meier analysis (Mizuno *et al.*, 2009) was carried out using the prognoscan database (http://www.prognoscan.org) to determine if the down-regulation of selected genes was associated with patients outcome such as, overall survival (OS), poor relapse-free survival (RPS) or disease free survival (DFS). OS refers to the time a patient survives after certain date of event such as surgery or a particular treatment with

or without any disease or disease recurrence (Chua, 2005). Similarly, DFS that is also called RFS, refers to the time a patient survives without having any disease or a disease recurrence (after the disease has been fully removed by a surgery and treated with adjuvant therapy) (Chua, 2005). DFS is generally calculated in every three years, which then serves as a secondary point for OS, which is calculated in five years durations if patients alive. As in other statistical test, $p<0.05$ was considered statistically significant. Prognoscan database utilizes cancer microarray data based on the published studies around the world mainly from Gene Expression Omnibus (GEO) datasets enabling to assess the risk of downregulation or upregulation of a gene in large number of patients (Mizuno *et al.*, 2009). Kaplan-Meier analysis was utilised for *BNC1*, *CCDC8* and *GALNT9* in independent datasets to identify if downregulation of these genes in primary breast cancer is associated with relapse free or disease free survival.

CHAPTER 4

DNA methylation analyses to identify metastatic suppressor genes dysregulated in BBM

4.1 Introduction

Formation of brain metastasis follows the multistep metastatic cascade (Hu *et al.*, 2009) in which cancer cells escape from the tumour-site, enter the surrounding tissues through invasion, flows through the lymph or blood vessels (intravasation), escape out of blood capillaries (extravasation), and finally invade into the brain and proliferate (Joyce & Pollard, 2009; Fidler *et al.*, 2010; Eichler *et al.*, 2011). Each of these stages is inefficient and rate limiting; only a minority of the tumour cells make it to the new destination. Therefore, successful metastatic outgrowth is a result of the cumulative ability of a cancer cell to find suitable unique microenvironment at each step in the metastatic cascade (Joyce & Pollard, 2009).

The spread of breast cancer to other distant organs takes place in an organized fashion first through the lymphatic system and then through the blood. The micrometastatic paradigm states that there may be an existence of micrometastatic lesions in breast cancer patients before the diagnosis of breast cancer that will later develop as a metastatic disease (Cianfrocca & Goldstein, 2004). It is known that the risk of BBM occurring early (<2 years after primary diagnosis) is associated with early onset tumours charaxterised by oestrogen receptor negative (ER-ve), human epidermal growth factor receptor 2 overexpression (HER2+ve) and triple negative (ER-ve/PR-ve/HER2-ve) phenotypes (van 't Veer *et al.*, 2002; Ryberg *et al.*, 2005; Sanna *et al.*, 2007; Metzger *et*

al., 2011). However, more than 50% of BBMs occur over five years after the primary tumour was diagnosed. Many of these late recurring brain metastases are derived from ER+ primary tumours (Kennecke *et al.*, 2007; Sperduto *et al.*, 2010; Zhang *et al.*, 2013).

Genomic alterations that provide the potential for metastatic growth can be characterised as either those that also drive primary tumour growth advantage, those that provide potential for dissemination and infiltration (such as those that drive EMT (Wang & Shang, 2013)) or those that enable continued growth within the microenvironment of the new organ (Joyce & Pollard, 2009). The common long lag-time between primary tumour diagnosis and recurrence of a detectable secondary tumour suggests that cells from these breast tumours undergo a period of dormancy (Aguirre-Ghiso, 2007; Lim *et al.*, 2012). These dormant cells are often found as micrometastases in bone marrow (Klein, 2003; Pantel *et al.*, 2008) and a number of genetic and epigenetic alterations acquired by breast tumour micrometastases of the bone have been characterised (Mundy, 2002; Park *et al.*, 2007). However, the presence of these micrometastases is not in itself a strong prognostic indicator for later metastatic disease (Klein, 2003; Pantel *et al.*, 2008). Moreover, very little is known about specific genomic alterations that facilitate the dissemination and colonisation of primary breast tumours in to the brain.

The most frequently metastasising tumours to brain are breast, lung, melanoma and RCC (Soffietti *et al.*, 2008). We selected genes that are frequently methylated in one of these tumour types. In addition, we selected genes downregulated in EMT that may contribute to BBM. Based on our hypothesis section 2.2.1 those genes dysregulated either in lung, RCC or melanoma may be commonly dysregulated in BBM. Therefore, we carried out a screen of candidate genes to investigate if these genes are frequently dysregulated

through promoter hypermethylation in BBM. It is likely that some genes, which are methylated in BBM are "passengers" in relation to metastasis and, as such, will also be methylated in primary breast tumours, contributing to primary breast tumorigenesis. Therefore, to identify genes, which are specifically dysregulated in breast tumours that metastasises to the brain we also determined the methylation status of these genes in a cohort of unrelated primary breast tumours and in matched pairs of breast tumours and BBM from the same patients (Figure 4.1). Therefore, we identified candidate genes that are either dysregulated early in tumour evolution (methylation is common to primary tumour and resulting BBM) or at a later stage only after the metastasising cells have left the breast and have probably reached to the brain (methylation occurs only in BBM).

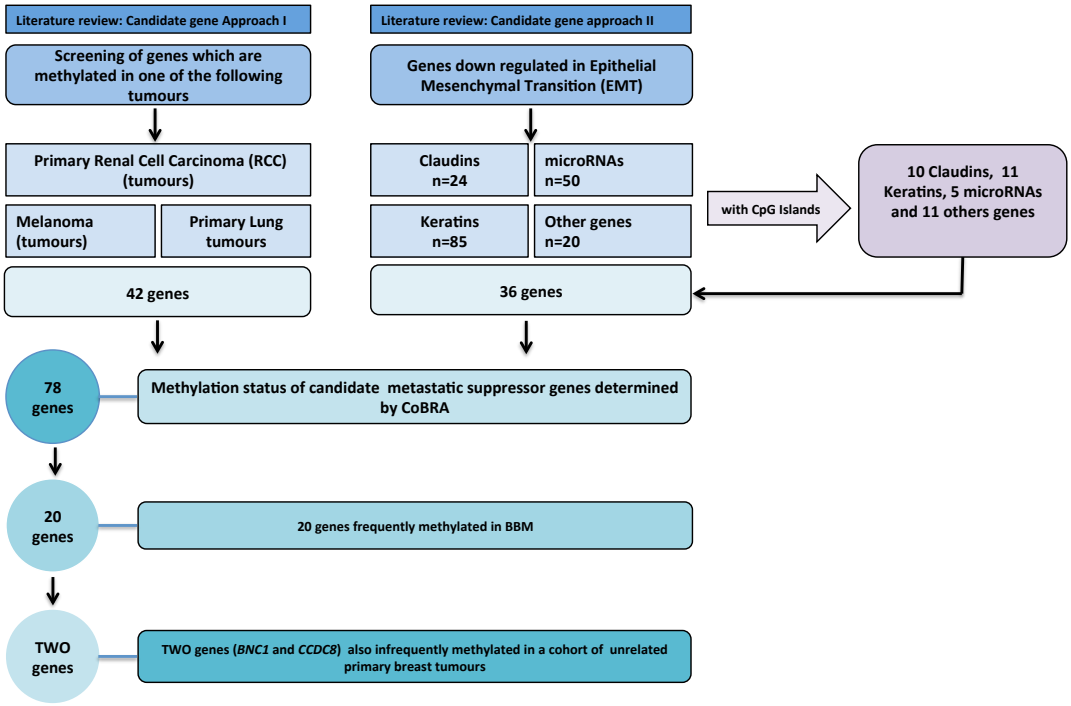


Figure 4.1: Summary of screening methodology and experimental analyses to identify genes dysregulated in BBM. Screening of genes from two independent approaches generated 78 candidate metastatic suppressor genes; Experimental analyses of the promoter regions of these 78 genes identified 20 genes that are frequently methylated in BBM samples, of which two genes were also infrequently methylated in a cohort of unrelated primary tumours.

4.2 Results

4.2.1 Screening of candidate metastatic suppressor genes in BBM

We have used two independent candidate gene approaches to identify genes dysregulated in breast tumours that metastasise to the brain. For the first approach, we carried out a comprehensive literature review to identify genes epigenetically dysregulated either in lung, breast, melanoma or renal tumours. This review generated independent lists of genes frequently methylated in lung, breast, melanoma and renal tumours. However, there were candidate genes that were commonly methylated in more than two or all tumours types including breast tumours. As we wished to identify genes uniquely dysregulated in BBM but not in primary breast tumours, we discarded the genes which were frequently methylated in primary breast tumours regardless of their methylation status in other tumour types. Careful screening of the long lists of genes generated 42 genes, which were methylated either in melanoma, lung or renal tumours. Moreover, we confirmed the methylation status of all these 42 genes in primary breast tumours using the 450K-methylation data from TCGA.

In a second approach, we included genes that have been downregulated in EMT regardless of their methylation status in primary tumours. However, we ensured that the genes identified are downregulated in EMT and have not been reported to be methylated in primary breast tumours. These candidate metastatic suppressor genes associated with EMT included claudins, keratins, tumour suppressor microRNAs, and other metastatic candidate genes. We identified 10 claudins, 9 keratins and 6 tumour suppressor microRNAs.

From both of these approaches, we generated a list of 78 genes with well-defined CpG islands. A full list of these genes and their biological function including associated references is given in appendix B1.

4.2.2 Identification of frequently methylated genes in metastatic brain tumours

We determined the methylation status of 78 candidate genes by CoBRA in up to 30 BBM samples. To ensure clinical significance, it is important that the genes we intend to identify are enriched in methylation status in the population of patients with BBM. Therefore, as an initial screening, 15 BBM tumours were analysed and those genes, which were methylated in less then 1/3rd of 15 BBM (5 of 15) samples were discarded. As we imposed a cut-off value for frequent methylation of a gene as ≥50% in BBM samples, it was unlikely that the genes which were methylated in less than 5/15 tumours in the first cohort, will be methylated sufficiently frequently in a second cohort to make up a frequency of ≥50% methylation in a total of 30 samples. Those genes that were methylated in more than 30% in the first cohort of 15 BBM samples were selected for methylation analysis in a second cohort of 20 unrelated primary breast tumours (section 4.2.3). Only those genes, which were infrequently methylated in this cohort of 20 primary tumours, were then investigated for their methylation status in a second cohort of 15 BBM samples.

From our panel of 42 candidate genes from the literature review, we identified 10 genes that were frequently methylated (>50%) in 15 BBM samples. These were *HOXD3* (100%), *CCDC8* (73%), *HOXB13* (80%), *ABCB1* (80%), *PENK* (80%), *BNC1* (68%), *PCDH8* (53%), *STAT3* (67%), *TNFRSF10D* (60%) and *WIF1* (53%) (table 4.1, (figure, 4.3, figure 4.4, appendix B2.). In addition, *SFRP2* was methylated in 47% of the BBM

samples. Therefore, we also considered screening its methylation in a cohort of 20 unrelated primary tumours.

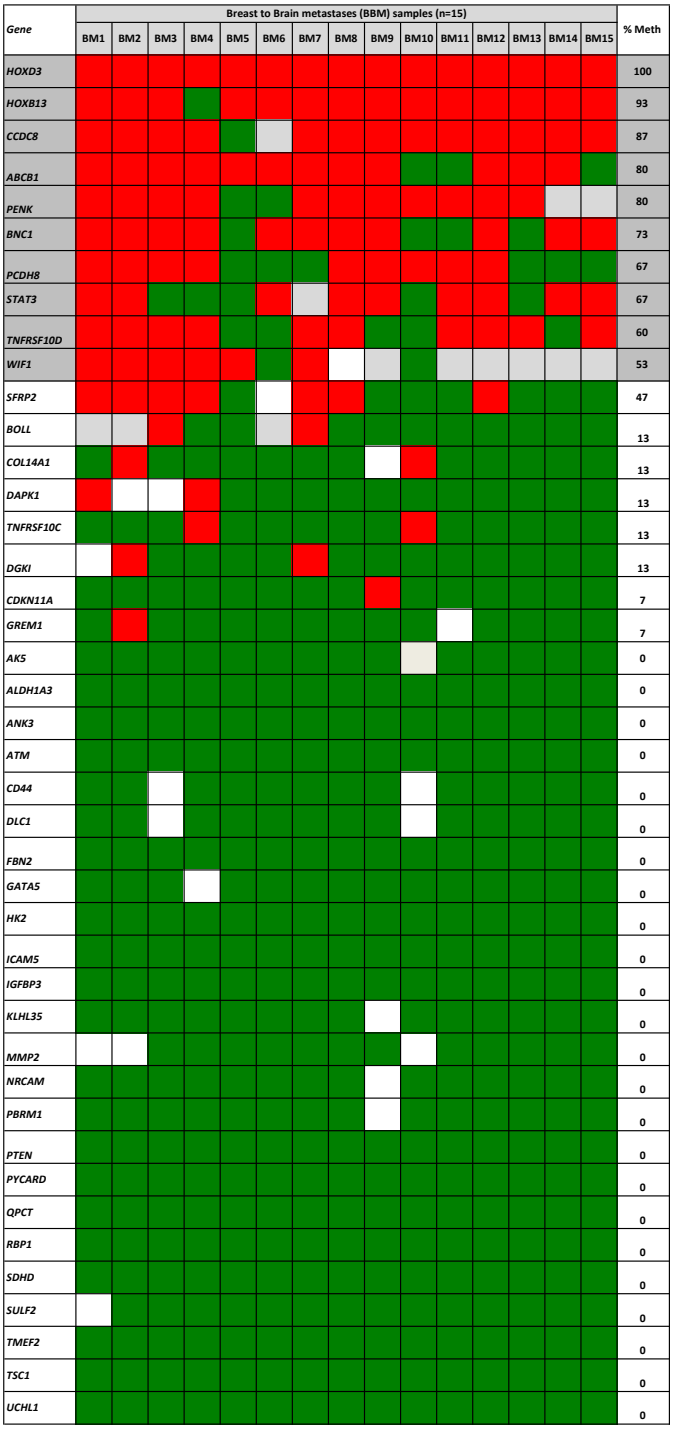


Table 4.1: Promoter methylation status of 42 candidate metastatic suppressor genes in BBM. Initially, a cohort of 15 BBM samples was used to investigate the methylation status of the genes. Ten genes (highlighted in gray) were frequently methylated (50% of the samples) in BBM. Red box indicates methylated samples, green box indicates

unmethylated samples and white box indicates samples that were not successfully amplified during CoBRA PCR.

In addition, from a panel of 36 genes downregulated in EMT, we identified 10 genes frequently methylated (>50%) in 15 BBM samples. These were *CLDN18* (100%), *KRT85* (100%), *MIR127* (100%), *MIR433* (100%), *MIR23b* (92%), *KRT83* (84%), *MST1R* (78%), *BVES* (64%), *CLDN6* (55%) and *HOXD10* (55%) (table 4.2). In addition, we also considered four other genes *CLDN5* (47%), *KRT72* (42%), *MIR124*-1 (36%) and *MIR34B* (33%) to analyse their methylation status in a cohort of unrelated primary breast tumours.

From both the panels, 20 genes were found to be frequently methylated in 15 BBM samples. The basic information (accession and full name) and the biological functions of these genes are given in table 4.3. We proceeded to determine the methylation status of these 20 genes and another five genes, with methylation frequency of ≥35%, (*SFRP2*, *CLDN5*, *KRT72*, *MIR124*-1 and *MIR34B*) in a cohort of unrelated 20 primary breast tumours.

Table 4.2: Methylation status of candidate metastatic suppressor genes in BBM

Gene	Breast to Brain metastases (BBM) samples (n=15)															% Meth (n=15)
	BM1	BM2	BM3	BM4	BM5	BM6	BM7	BM8	BM9	BM10	BM11	BM12	BM13	BM14	BM15	
CLDN18																100
KRT85																100
MIR127																100
MIR433																100
MIR23B																92
KRT83																84
MST1R																78
BVES																64
CLDN6																55
HOXD10																55
CLDN5																47
KRT72																42
MIR124-1																36
MIR34b																33
CLDN1																20
MIR34a																10
CLDN3																0
CLDN4																0
CLDN7																0
CLDN9																0
CLDN11																0
CLDN23																0
CSNK1A1																0
CMTM8																0
DSP																0
FBXL14																0
KRT7																0
KRT18																0
KRT19																0
KRT28																0
KRT81																0
OCLN																0
PNN																0
TJP1																0
TSPAN13																0

Table 4.2: Promoter methylation status of 36 candidate genes, which are down regulated in Epithelial Mesenchymal Transition (EMT). Initially, a cohort of 15 BBM samples was used to investigate the methylation status of the genes. Ten genes (shaded in grey) were frequently methylated (50% of the samples) in BBM. Red box indicates methylated samples, green box indicates unmethylated samples and white box indicates samples that were not successfully amplified during CoBRA PCR.

Table 4.3: Genes frequently methylated in BBM samples and their functions.

Gene symbol	Accession	Gene name	% of tumours methylated	Function
CLDN18	NM_016369.3	*Claudin 18*	100	Intercellular adhesion molecule responsible for tight junction strand formation (Tsukita & Furuse, 2000)
KRT85	NM_002283.3	*Keratin 85*	100	Component of intermediate filament in epithelial cells contributing to cell-cell adhesion (Magin *et al.*, 2007; Bragulla & Homberger, 2009; Shimomura *et al.*, 2010)
MIR127	NR_029696.1	*microRNA 127*	100	Regulator of cell proliferation and senescence (Chen *et al.*, 2013)
MIR433	NR_029966.1	*microRNA 433*	100	Deregulated in gastric cancer, regulator of cell migration and drug response (Luo *et al.*, 2009; Symmans, 2010)
HOXD3	NM_006898.4	*HomeoboxD3*	100	Proangiogenic transcription factor (Chen *et al.*, 2004)
MIR23B	NR_029664.1	*microRNA 23b*	92	Involved in cytoskeleton modelling, motility and metastasis (Majid *et al.*, 2012; Zaman *et al.*, 2012; Jin *et al.*, 2013; Pellegrino *et al.*, 2013)
CCDC8	NM_032040.4	*Coil coiled domain containing 8*	87	Mutated in patients with 3M syndrome (Hanson *et al.*, 2011). Loss is associated with

				genomic instability and aneuploidy (Yan *et al.*, 2014).
KRT83	NM_002282.3	*Keratin 83*	84	Component of intermediate filament, contributes to cell to cell adhesion (Magin *et al.*, 2007; Bragulla & Homberger, 2009)
HOXB13	NM_006361.5	*Homeobox B13*	80	TSG for prostate cancer, inhibits androgen mediated signalling (Fidler *et al.*, 2010)
ABCB1	NM_000927.4	*ATP-binding cassette sub-family B member 1*	80	Controls efflux of substances across plasma membranes, associated with multidrug resistance (Muggerud *et al.*, 2010)
PENK	NM_006211.3	*Proenkephalin*	80	Promotes RNA splicing in osteoblasts and neural cells, plays role in bone development (Rosen *et al.*, 2013b)
MST1R	NM_002447.2	*macrophage stimulating 1 receptor*	78	Involved in intracellular signalling cascades leading to cellular growth, motility and invasion (Wagh *et al.*, 2008)
BNC1	NM_001717.3	*Basonuclin 1*	73	Zink finger transcription factor, regulator of EMT (Feuerborn *et al.*, 2014)
PCDH8	NM_002590.3	*Procadhern 8*	73	Helps in cell to cell adhesion (Sabine *et al.*, 1998)
STAT3	NM_139276.2	*Signal*	67	Involved in embryonic

		transducer and activator of transcription 3		stem cell regulation, somatic cell growth (KIYOSHI *et al.*, 1997; Hitoshi *et al.*, 1998; Akira)
BVES	NM_007073.4	*Blood vessel epicardial substance*	64	Involved in inter-cellular interaction and cell adhesion. (Osler *et al.*, 2006)
TNFRSF10D	NM_003840.4	*Tumour Necrosis Factor receptor superfamily 10 D*	60	Member of TNF(Tumour Necrosis Factor) receptor superfamily, promotes apoptosis in cancer cells (Hill *et al.*, 2011)
CLDN6	NM_021195.4	*Claudin 6*	55	Intercellular adhesion molecules responsible for tight junction strand formation, its epigenetic silencing is associated with migration and invasiveness of breast cancer (Tsukita & Furuse, 2000; Osanai *et al.*, 2007)
HOXD10	NM_002148.3	*Homeobox D10*	55	Maintain epithelial cell plasticity and contributes to stability of extracellular matrix (Carrio *et al.*, 2005)
WIF1	NM_007191.4	*Wnt inhibitory factor-1 gene*	53	Inhibitor of Wnt-signalling (Ai *et al.*, 2006; Veeck *et al.*, 2009)

Table 4.3: Genes frequently methylated in breast to brain metastases. 20 genes are frequently methylated in brain metastases (n=15). These genes were further analysed in 20 primary breast samples (n=40 in total) and 15 breast to brain metastases (n=30 in total).

4.2.3 Determination of methylation status of 20 genes in a cohort of unrelated primary breast tumours

We screened primary breast tumours for promoter methylation status of the genes that were frequently methylated (>30%) in BBM. To ensure that genes identified in this study are clinically significant, we imposed a relatively low cut off frequency of ≤45% for methylation in primary breast tumours. This fits our hypothesis that genes deregulated in the process of BBM will either occur early in tumour evolution and be detectable in the primary tumours that eventually metastasise (BBM occurs in 18-30% of breast tumours (Weil *et al.*, 2005; Gori *et al.*, 2007; Kennecke *et al.*, 2007; Tosoni *et al.*, 2008)) or it occurs later in tumour evolution and will not be detectable in the primary tumour. The clinical information of unrelated primary tumours used for this analysis is given in appendix B4.

From a panel of 10 genes frequently methylated in brain metastases (from our literature review candidates), we identified that 8 of these genes are also frequently methylated in primary breast tumours. These are *HOXD3 (81%), HOXB13 (53%), ABCB1 (68%), PCDH8 (54%), PENK (79%), STAT3 (57%) TNFRSF10D (75%)* and *WIFI (55%)* (table 4.4, appendix B5). In addition, *SFRP2*, which was methylated in 47% of the BBM samples, was methylated in 90% of the primary breast tumour samples (appendix B5). This suggests that *SFRP2* is hypomethylated in BBM compared to unrelated primary breast tumours and also frequently methylated in both the primary tumours and BBM. This suggests that these genes are not uniquely epigenetically dysregulated during the process of BBM. Only two genes, *BNC1* (21%) and *CCDC8* (40%) were infrequently methylated in these unrelated primary tumours. Therefore, these two genes were further

investigated for their methylation status in a further 20 unrelated primary tumours (40 primary tumours in total). It is worth noting that to our knowledge this is the first time that promoter methylation in *CCDC8*, *HOXD3, PCDH8*, *PENK, STAT3, SFRP2 and WIF1* has been described in primary breast tumours.

Promoter methylation of *BNC1* (21%) and *CCDC8 (*40%) in primary breast tumours was infrequent (≤45%), and was statistically significantly lower than the frequency of methylation in BBM (p=0.0001, p=0.01 respectively) (figure 4.2, figure 4.3 and figure 4.4).

From a panel of 10 genes down regulated in EMT that are frequently methylated in brain metastases, we found all 10 genes also to be frequently methylated in primary breast tumours *i.e. CLDN18* (100%), *KRT85* (100%), *MIR127* (100%), *MIR433* (100%), *KRT83* (100%), *MIR23b* (60%), *MST1R* (60%), *BVES* (60%), *CLDN6* (50%), *HOXD10* (55%) (table 4.4). The high frequency of methylation in primary tumours indicates that epigenetic deregulation of these genes is not driving BBM. This is the first time that promoter methylation has been described for all of these 10 genes.

In addition, *MIR34B* is also methylated in 60% of the primary tumours. However, it was methylated in 33% of the BBM samples. *KRT72* was methylated in 25% of the primary tumours, which was lower than its methylation in BBM (42%), However this was not statistically significant (p=0.26). The other two genes *CLDN5*, and *MIR124*-1 were methylated in 17% and 19% of the primary tumours respectively. Due to their relatively infrequent methylation in primary tumours compared to BBM (47% and 36% respectively), we also analysed these genes in a second cohort of 15 BBM to investigate if their differential methylation statuses are statistically significant.

Table 4.4: methylation status of candidate metastatic suppressor genes in primary breast tumours

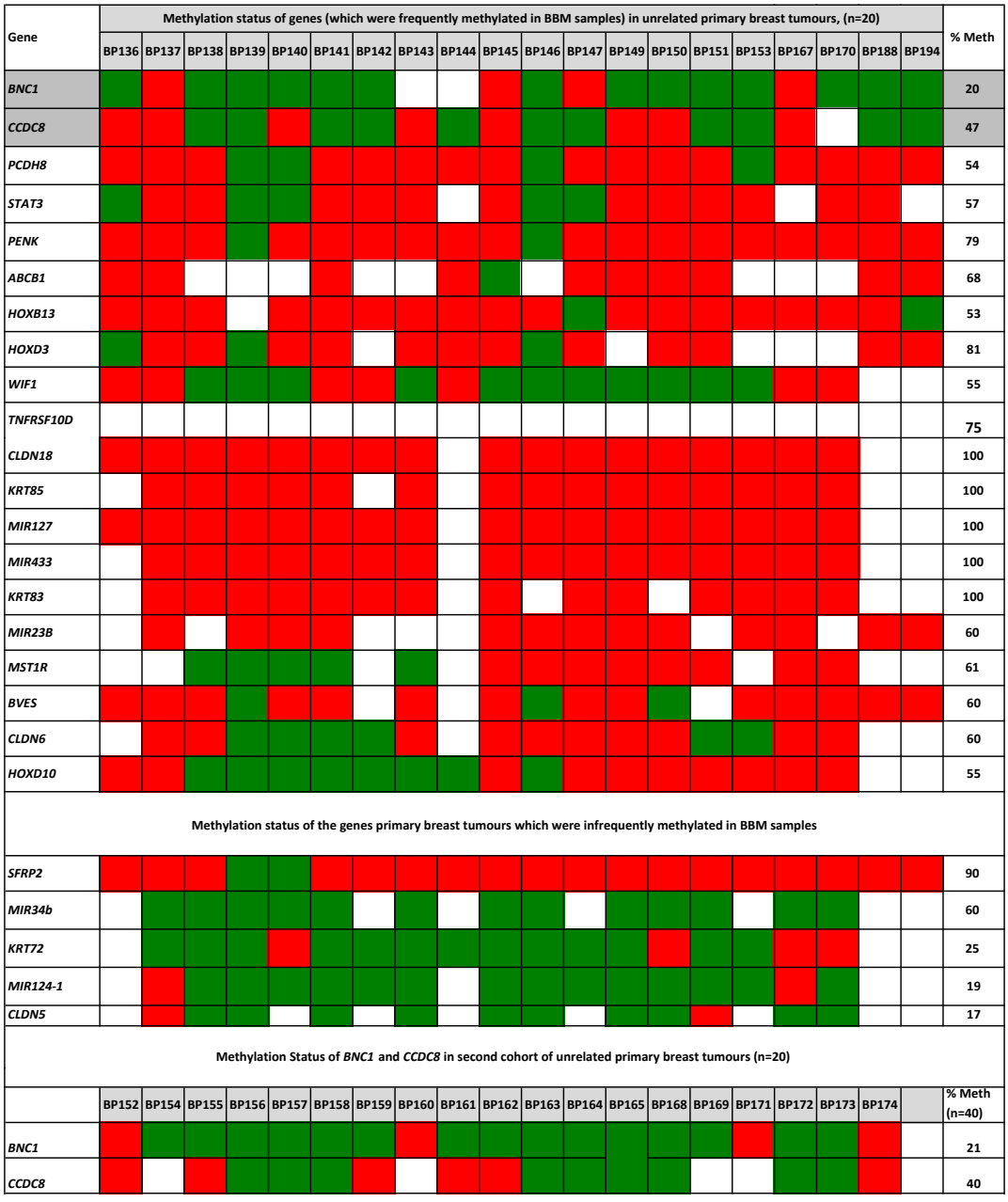


Table 4.4: Promoter methylation status of 20 candidate metastatic suppressor genes, which were frequently methylated in BBM samples, were analysed in an unrelated cohort of 20 primary breast tumours. Five other genes were also analysed in primary breast tumours to investigate their methylation status (section 4.2.3). Only two genes; *BNC1* and *CCDC8* are differentially methylated in BBM (frequently methylated) and primary breast tumours (infrequently methylated). They were

further analysed in a second cohort of 20 unrelated primary breast tumours, making up to a total of 40 primary tumours. The red box indicates methylated samples, green box indicates unmethylated samples and white box indicates samples that were not successfully amplified during CoBRA PCR.

The four genes; *BNC1, CCDC8, CLDN5* and *MIR124*-1 were further analysed in a second cohort of 15 BBM samples (table 4.5). *CLDN5* and *MIR124*-1 were analysed in a second cohort of 15 BBM samples due to their differences in the methylation in BBM and primary tumours. The aim of this project was to identify genes, which were frequently methylated in BBM and infrequently methylated in primary tumours. The methylation frequency of *CLDN5* and *MIR124*-1 is higher in BBM (40% and 36% respectively) than in primary breast tumours (17% and 19% respectively), however they were infrequently methylated both in primary breast tumours and BBM. The genes are enriched in methylation in BBM, but this is not statistically significant (p=0.07 and 0.17 respectively); therefore, we excluded these genes from further investigation.

Taken together, from the list of 78 candidate metastatic suppressor genes based on our broad-ranging screen of BBM candidate genes, only two genes with significantly differing methylation status in primary breast tumours and BBM were identified. The significant difference in frequency of methylation of *BNC1* (figure 4.3) and *CCDC8* (figure 4.4) in primary breast tumours and BBM may be an indication that these genes may contribute to BBM and are good candidates for further investigation.

Gene	BM16	BM17	BM18	BM19	BM20	BM21	BM22	BM23	BM24	BM25	BM26	BM27	BM28	BM29	BM30	BM31	% meth (n=31)
BNC1																	68
CCDC8																	73
CLDN5																	40
MIR124-1																	36

Table 4.5: Promoter methylation status of candidate metastatic suppressor genes in a second cohort of 16 BBM samples. Two of these genes *BNC1* and *CCDC8* are frequently methylated in 31 BBM samples and infrequently methylated in 40 unrelated primary breast tumours. The methylation frequency of *CLDN5* and *MIR124*-2 was higher in the initial cohort of 15 BBM samples. This analysis confirmed that *CLDN5* and *MIR124*-2 are infrequently methylated in both BBM and primary tumours. However, they are enriched in methylation in BBM compared to primary breast tumours. Red box indicates methylated samples, green box indicates unmethylated samples and white box indicates samples that were not successfully amplified during CoBRA PCR.

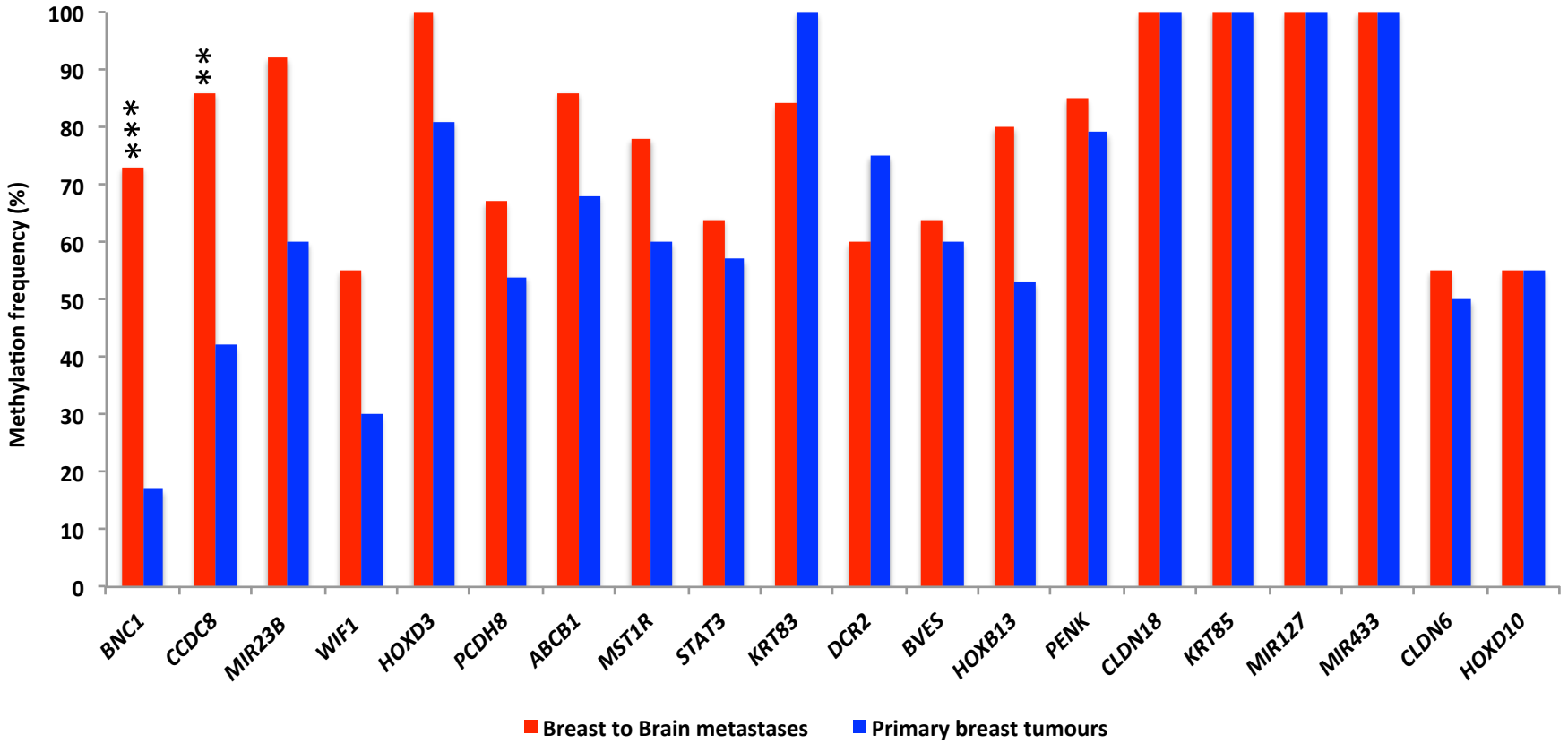


Figure 4.2: Methylation frequency of candidate metastatic suppressor genes in breast-to-brain metastases (BBM) (n=30) versus primary breast tumours (n=40). Out of 20 genes that are frequently methylated in brain metastases, two genes (*BNC1* and *CCDC8*) are infrequently methylated in a cohort of independent primary tumours with statistical significance (*: p≤ 0.0001 and **: p ≤0.01).**

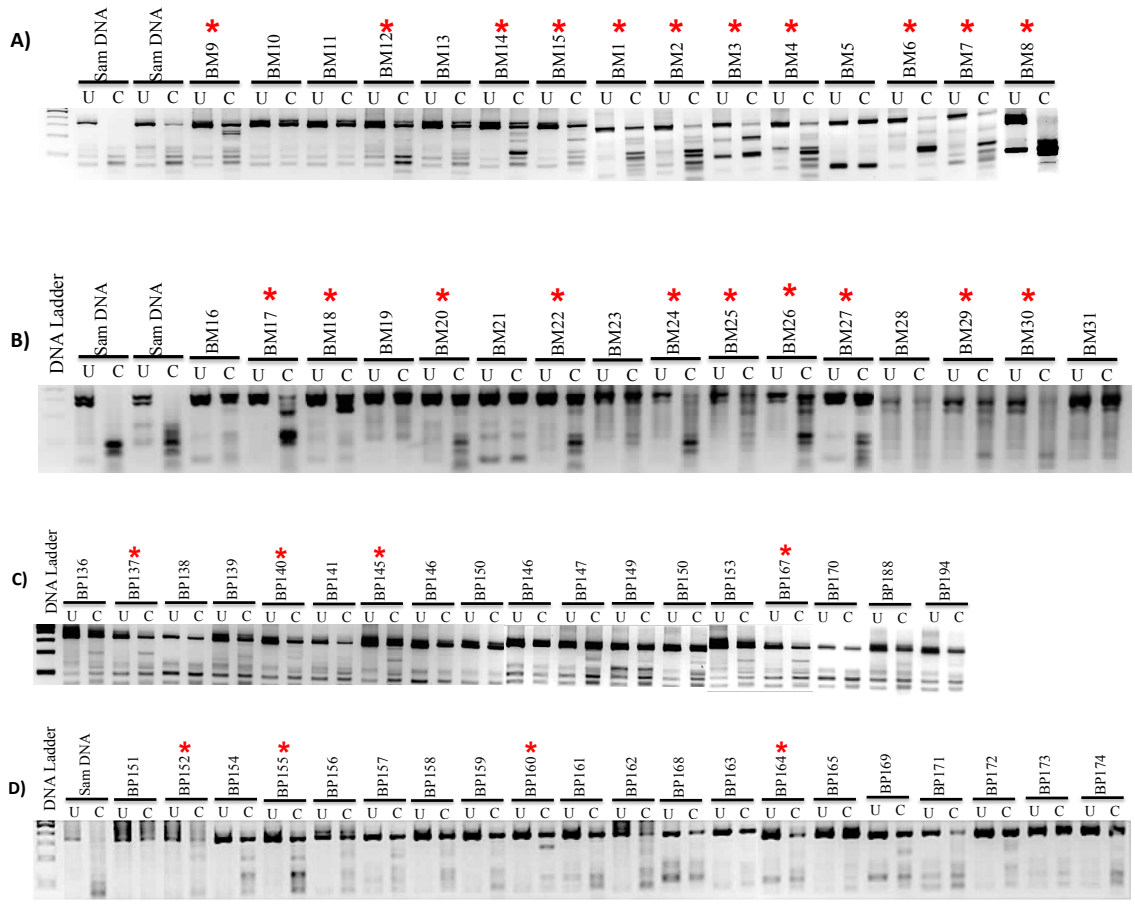


Figure 4.3: Methylation status of *BNC1* in the first and second cohort of BBM samples (A and B respectively) and in the first and second cohort of unrelated primary breast tumours (C and D respectively). *BNC1* is frequently methylated (68%) in metastatic brain tumours and is infrequently methylated (21%) in primary breast tumours. SAM: Fully methylated positive control, BM: Brain metastases, BP: primary breast tumours U: Uncut/control sample, C: cut by restriction enzyme, *: methylated samples

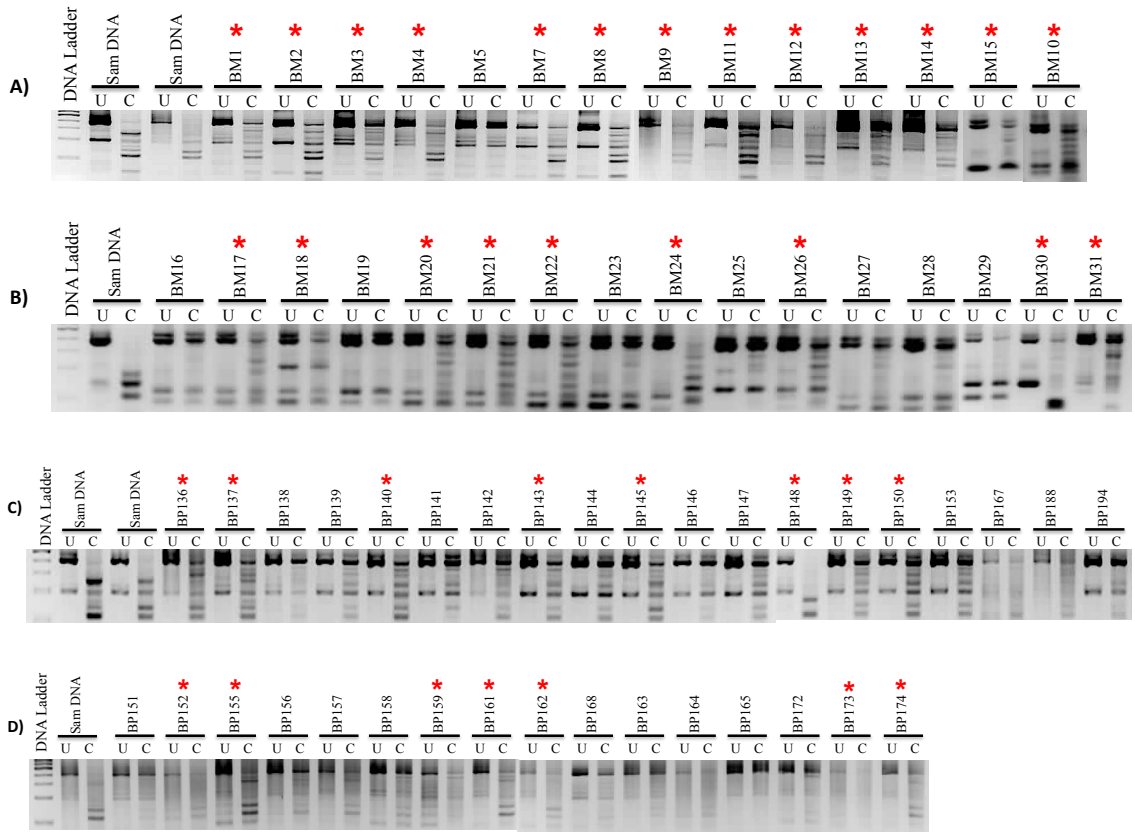


Figure 4.4: Methylation status of *CCDC8* in the first and second cohort of BBM samples (A and B respectively) and in first and second cohort of unrelated primary breast tumours (C and D respectively). *CCDC8* is frequently methylated (73%) in BBM (n=30) and is infrequently methylated (40%) in primary breast tumours (n=35). SAM: Fully methylated positive control, BM: Brain metastases, BP: Primary breast tumours, U: Uncut/control sample, C: cut by restriction enzyme, *: methylated samples.

4.2.4. Expression analysis of *BNC1* and *CCDC8* in metastatic brain tumours

Having identified two candidate genes that are differentially methylated in primary breast tumours and metastatic brain tumours we proceeded to determine if this promoter methylation correlated to gene expression.

Total RNA was extracted from 15 metastatic brain tumours to determine the expression of *BNC1* and *CCDC8,* by RT-PCR. *BNC1* and *CCDC8* were frequently downregulated or silenced in these tumours and reduced expression correlated to promoter methylation (Figure 4.5). Of 15 metastatic brain tumours, *BNC1* was methylated and silenced or downregulated in 10 tumours. *BNC1* was expressed in two unmethylated samples (BM11 and BM23); downregulated in two partially methylated samples (BM12 and BM18) and are silenced in seven methylated samples (BM13, BM24, BM17, BM20, BM27, BM29 and BM30). *BNC1* is also silenced in three unmethylated tumours (BM13, BM16 and BM28). Similarly, out of 15 available tumours for expression analysis, *CCDC8* was methylated and silenced or downregulated in 10 samples. *CCDC8* is expressed in four unmethylated samples (BM16, BM23, BM28 and BM29); down regulated in two partially methylated samples (BM20 and BM30) and silenced in 6 methylated samples (BM3, BM11, BM12, BM14, BM15 and BM18). *CCDC8* is also expressed in two methylated samples (BM13 and BM17). *CCDC8* is silenced in one unmethylated sample (BM27).

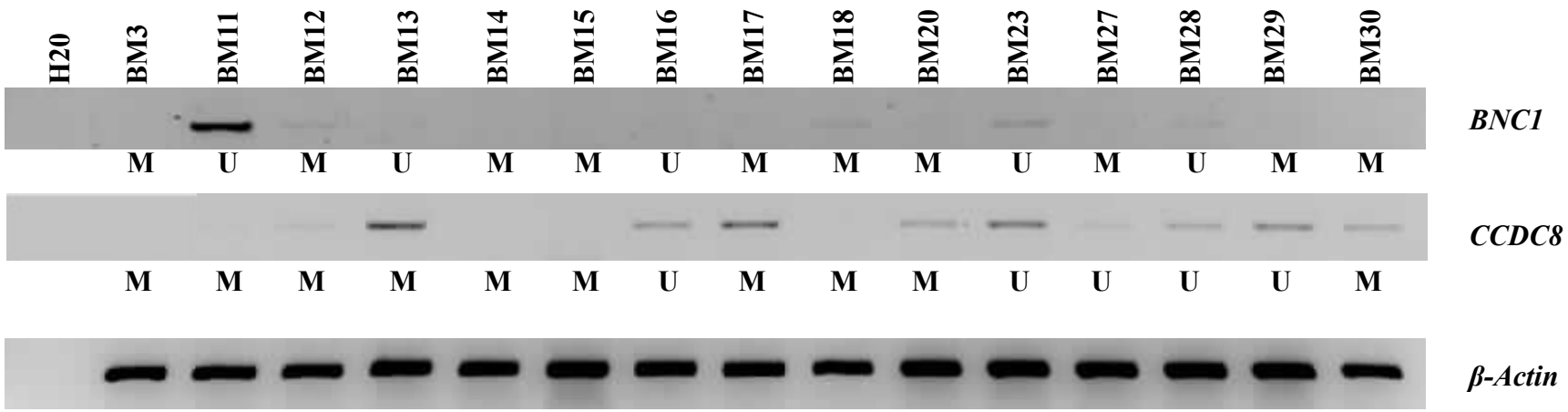


Figure 4.5: Gel electrophoresis of reverse transcription PCR (RT-PCR) of *BNC1* and *CCDC8* in BBM. Expression analysis using RT-PCR shows that the expression of *BNC1* and *CCDC8* correlates with its promoter methylation in BBM. *BNC1* and *CCDC8* are expressed in unmethylated (U), down regulated in partially methylated (*), and is silenced in fully methylated (**) tumours. (See figure 4.3 and 4.4 for methylation analysis). Expression of β-actin was determined to ensure equal loading for all samples. Of 15 BBM samples available for expression analysis, *BNC1* is expressed in two unmethylated samples (BM11 and BM23); downregulated in two partially methylated samples (BM12 and BM18) and are silenced in six methylated samples (BM3, BM17, BM20, BM27, BM29 and BM30). *BNC1* is also silenced in three unmethylated tumours (BM13, BM16 and BM28). Similarly, *CCDC8* is expressed in four unmethylated samples (BM16, BM23, BM28 and BM29); down regulated in two partially methylated samples (BM20 and BM30) and are silenced in six methylated samples (BM3, BM11, BM12, BM14, BM15 and BM18). *CCDC8* is also expressed in two methylated samples (BM13 and BM17). CCDC8 is silenced in one unmethylated sample (BM27). M: Methylated samples, U: Unmethylated samples.

4.2.5 Promoter methylation status of *BNC1* and *CCDC8* in brain metastases and associated primary breast tumours from individual patients

We analysed the methylation status of two genes *BNC1* and *CCDC8* in matched pairs of tumours *i.e.* metastatic brain tumours and corresponding primary tumours from individual patients. We have 10 of these pairs, however, some loci in the primary tumour DNA proved refractive to amplification. Out of 8 matched pairs, where the *BNC1* promoter region was successfully amplified, the region is methylated in all 8 of the brain metastases. However, it is only methylated in one corresponding primary tumour (Figure 4.6A). In contrast, out of 11 matched pairs, *CCDC8* is commonly methylated in 10 corresponding primary tumours (Figure 4.6B).

These results suggest that *BNC1* promoter methylation occurs at a late stage in the evolution of metastatic brain tumours, possibly after they have metastasised to the brain. Alternatively, methylation of these genes may occur in a small subset of cells within the primary tumour (below the detection threshold of this assay) and these cells are enriched in the metastatic tumour. In contrast, *CCDC8* promoter methylation is detectable in most primary tumours that metastasise to the brain, suggesting that it may play an important role in the early stages of primary tumour metastasis.

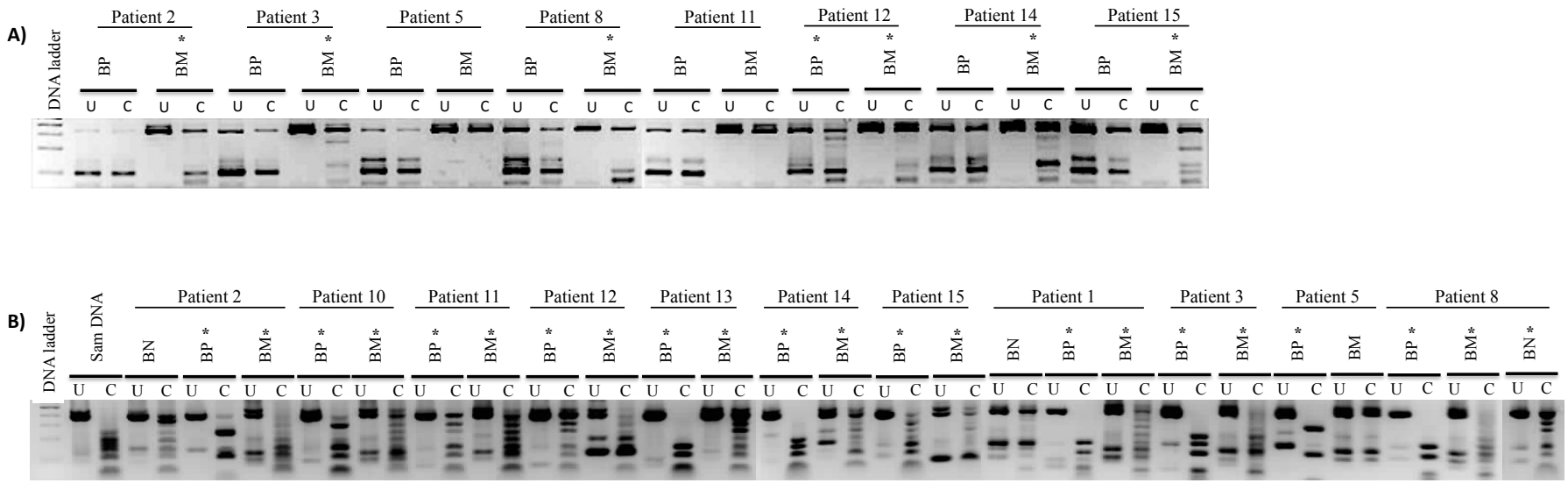


Figure 4.6: Methylation status of *BNC1* and *CCDC8* in metastatic brain tumours and their corresponding originating primary breast tumours from individual patients. CoBRA was used to determine the methylation status; small, digested PCR products in the BstuI cut (C) lane compared to the undigested (U) lane indicates promoter methylation in a sample. *BNC1* was frequently methylated (*) in metastatic brain tumours (BM) and NOT methylated in the originating breast primary (BP) tumours. Of 8 matched pairs analysed, *BNC1* is methylated in 6 metastatic brain tumours (see patient 2, 3, 8, 12 and 15), whereas it is methylated in only one of the corresponding primary tumour (patient 12). Similarly, *CCDC8* was frequently methylated (*) in metastatic brain tumours (BM) and their originating breast primary (BP) tumours. Of 11 matched pairs analysed, *CCDC8* is methylated in 10 metastatic brain tumours (see patient 2, 10, 11, 12, 13, 14, 15, 1, 2, and 8) and all 11 corresponding primary tumours. Interestingly,

***CCDC8* is also methylated in one primary tumour (patient 5) where its corresponding BBM is not methylated. BP: Breast Primary tumour, BM: Metastatic Brain tumour, BN: adjacent Normal Breast tissue, U: Uncut/Control sample, C: cut by methylation specific restriction enzyme, *: Methylated samples, Sam DNA: fully methylated positive control.**

4.3 Discussion:

Metastasis to the brain is an increasingly common event in the progression of breast cancer (Sperduto *et al.*, 2010). This trend is likely to continue as primary tumour management improves. Given the extremely poor clinical outcome following a diagnosis of BBM (Sperduto *et al.*, 2010). it is imperative that the underlying molecular biology that drives tumour evolution to the colonization of the brain is revealed. An improved understanding of these events will identify novel therapeutic targets and prognostic markers.

To date some progress has been made to identify prognostic markers for breast cancer metastasis by gene expression profiling (Paik *et al.*, 2004). However, prediction of site specific-metastasis remains poor (Weigelt *et al.*, 2005).

The importance of gene dysregulation by promoter methylation as a mechanism of tumour evolution is now well established (Cock-Rada & Weitzman, 2013). Indeed, genome wide methylation analysis of many hundreds of primary breast tumours has allowed the definition of specific sub-categories of breast tumours (Curtis *et al.*, 2012; TCGA, 2012) and our increasing understanding of the molecular basis of these subtypes has improved our ability to predict early metastatic recurrence (Metzger *et al.*, 2011; Stephens *et al.*, 2012). However, late recurrence, a common feature of BBM or indeed any site specific recurrence has proven difficult to predict (Burstein & Griggs, 2012).

We carried out a broad candidate approach to identify genes that are dysregulated in BBM. This approach consisted of identifying two separated candidate gene lists. The first cohort of candidate genes was chosen after considering the observation that the period between diagnosis of primary breast tumour and brain metastasis is often very

much longer than that of lung, renal or melanoma tumours (Eichler *et al.*, 2011). We hypothesized that epigenetic events that are common in lung, renal and melanoma tumours and drive early tumour development may also contribute to brain metastasis. However, these changes may not contribute to primary breast tumour formation but do have the capacity to contribute to BBM. Thus, the genes that contribute to BBM and are infrequently methylated in primary breast tumours (and commonly methylated in other primary tumours) will be frequently methylated in BBM. We carried out a literature review for genes methylated in these tumour types and selected only those genes that were also infrequently methylated in primary breast tumours according to Illumina HumanMethylation 450k data from *TCGA*. From our resulting candidate list of 42 genes, we identified *BNC1* and *CCDC8* as differentially methylated in primary breast tumours and BBM, and the methylation of these genes was associated with silencing.

In addition, we screened a selection of genes that have previously been shown to be dysregulated during the process of epithelial to mesenchymal transition, a well-established mechanism during the process of metastasis (Thiery, 2002; Yang & Weinberg, 2008). However, following analysis of these genes, we found no significant difference between the promoter methylation status of those genes in BBM or primary breast tumours; genes found to be frequently methylated in BBM were also frequently methylated in primary breast tumours. This suggests that none of these frequently methylated genes are involved in the specific process of BBM. However, it is possible that frequently methylated genes may be involved in the process of breast cancer metastasis to other sites such as bone, lungs or liver. Moreover, these genes could possibly be involved in the formation of micrometastases (Hanssens *et al.*, 2011) in other body sites such as bone or lung and, other additional genetic and epigenetic

alterations may further contribute to those micrometastases to metastasise into the brain. Therefore, further analyses may be useful to understand the role of these genes in the process of breast cancer metastases and BBM. Therefore, these genes could possibly be involved both in primary tumours and BBMs. Cells that disseminate from primary tumours through the process of EMT often revert back to an epithelial phenotype once they have infiltrated the site of metastasis through a process of Mesenchymal to epithelial transition (MET) (Yang & Weinberg, 2008). It will require subtle experimental models to track such changes through the process of tumour evolution.

We identified 42 genes that had previously been reported to be frequently methylated in either Lung, renal, or melanoma tumours but 450K methylation array analysis (using data from TCGA) showed these genes were infrequently methylated in primary breast tumours that had, at the time of diagnosis, shown no indication of metastasis. However, our independent analysis of a second cohort of 40 unrelated primary breast tumours revealed that, contrary to our analysis of genome-wide array data, many of these genes were frequently methylated in both BBM and the primary breast tumours. This analysis identified two genes (*BNC1* and *CCDC8*,) that are differentially methylated in primary breast tumours and BBM and merit further investigation.

We predicted that our analysis of unrelated primary breast tumours and BBM would identify two different classes of genes that contribute to BBM. We hypothesised that epigenetic silencing of BBM associated genes would either occur as a) early events in tumour evolution that may be involved in processes such as local invasion and intravasion (Chambers *et al.*, 2002; Gupta & Massague, 2006) or these early events may be required for specific distant site metastasis but also contribute to primary tumour

development. Or, b) late events that play no significant role in the initial evolution of the primary tumour but contribute to the development of the secondary brain metastasis, either by promoting invasion or improving the capacity for these foreign cells to survive in the novel microenvironment of the brain.

The existence of early and late events has previously been proposed by Nguyen *et. al.* (Nguyen *et al.*, 2009), they classified deregulated genes as either involved in: (i) Metastasis initiation (*e.g.* EMT regulators), these will be detectable in the primary tumour. (ii) Metastasis progression genes that are important for survival in the circulation or required for extravasation, these may be detectable in the primary tumour, however they may occur once metastasising cells have left the primary site, or (iii) metastasis virulence genes that allow the cancer cells to survive in a foreign tissue environment. These are likely to occur as a consequence of the selection pressure provided by the novel environment the metastasised tumour cells find themselves in and as such will be a late event in metastasis evolution. Metastasis progression genes may have different functions in the primary tumour and distant metastasis, for example *MMP-1* promotes vascular re-modelling in primary breast tumours and also contributes to lung extravasation (Minn *et al.*, 2005a). An example of known metastasis virulence genes that does not contribute to primary tumour growth is interleukin-11, which promotes breast tumour metastasis to the bone but does not provide any advantage to the primary tumour (Kang *et al.*, 2003).

Both early and late methylation events will appear similarly in our initial analysis; the genes will be frequently methylated in BBM and infrequently methylated in unrelated primary breast tumours. However a comparison of primary tumours and BBM from the

same patient should reveal if specific gene methylation occurs early or late in the process of tumour evolution.

BNC1 is mainly expressed in basal keratinocytes, squamous epithelium and in reproductive cells (Yang *et al*., 1997). BNC1 is known to function as a transcription factor in the synthesis of ribosomal RNA by interacting with rRNA gene promoter (Zhang *et al*., 2007). It is a zinc finger transcription factor that interacts with the promoters of both RNA polymerases I and II (Zhang *et al.*, 2007). BNC1 target genes have been implicated in a broad range of functions including chromatin structure, transcription/ DNA-binding, adhesion, signal transduction, and intracellular transport (Ma *et al.*, 2006; Wang *et al.*, 2006; Zhang *et al.*, 2007). It is expressed in a broad range of tissue types and is highly expressed in the testis (Safran *et al.*, 2010).

BNC1 has previously been shown to be silenced by promoter methylation in lung (Shames *et al.*, 2006b) renal, (Morris *et al.*, 2010), pancreatic (Yi *et al.*, 2013), prostate (Devaney *et al.*, 2013) and leukemic cancers (Dunwell *et al.*, 2009a). *In vitro* assays have shown that loss of BNC1 expression is associated with an increased malignant phenotype (Morris *et al.*, 2010). Consistent with this study, analysis of HumanMethylation 27K and 450K array data from *The Cancer Genome Atlas* indicates that *BNC1* Promoter methylation is an infrequent event in primary breast tumours TCGA (2012). However, frequent *BNC1* promoter methylation (>60%) in a small cohort of breast tumours has previously been reported (Shames *et al.*, 2006b) Cancer-associated *BNC1* mutations or copy number changes are rare (TCGA, 2012).

Expression of BNC1 is induced by transforming growth factor-β1 signalling and, in turn, it acts as a transcription factor for a number of modulators of epithelial dedifferentiation during the process of EMT (Feuerborn *et al.*, 2014). Moreover, loss of BNC1

expression results in a reduced EMT phenotype. These findings suggest that the expression of BNC1 would enhance the process of metastasis via EMT. Our findings are consistent with this; we find that *BNC1* is infrequently methylated in primary breast tumours (17%) and frequently methylated and silenced in BBMs (73%). Moreover, we have shown that *BNC1* promoter methylation is a late event in tumour evolution, only occurring in the brain metastasis of a BBM patient and not in the associated primary tumour. It is plausible that BNC1 expression is commonly required for EMT to occur during metastasis and, once these cells have metastasised to the brain, loss of *BNC1* expression contributes to mesenchymal to epithelial transition (MET).

An *in vitro* screen that consisted of multiple rounds of breast cancer cell line injection into nude mice and re-culturing of the resulting brain metastases showed that *BNC1* was among 373 genes that were differentially expressed (Guo *et al.*, 2011). However, contrary to the results presented here, BNC1 was overexpressed in the cells recovered from the mouse brain metastases. This apparent difference in expression may be as a consequence of the model used, or the specific expression profile of the cell line used (GI-101A). Alternatively, it may represent important differences in the process of aggressive early metastasis (as nude mouse cell line injection models represent) and slower metastatic evolution, where tumour cells proceed through a phase of latency or micrometastasis. As is often the case for brain metastases removed by surgery (Sperduto *et al.*, 2010; Jenkinson *et al.*, 2011), many of the brain metastases in our study were identified several years after initial breast cancer diagnosis (Paired primary and BBM samples were excised between 2-10 years apart).

CCDC8 encodes a coiled-coil domain containing protein (CCDC8) that is one of three proteins that are, mutually exclusively, mutated in patients with 3M syndrome (Hanson

et al., 2011). 3M syndrome is an autosomal recessive disorder characterised by short stature, skeletal abnormalities, reduced male hormone and blood vessel bulges (Huber *et al.*, 2005; Maksimova *et al.*, 2007; Huber *et al.*, 2009) *CCDC8* is mutated in ~5% of 3M cases, the other genes, *CUL7* and *OBSL1* are mutated in ~65% and ~30% of cases respectively (Hanson *et al.*, 2011; Huber *et al.*, 2011). It has recently been shown that these three proteins form a complex (the 3M complex) and that loss of expression of any one protein disrupts microtubule dynamics resulting in dysregulated mitosis and cytokinesis and associated genomic instability and aneuploidy (Yan *et al.*, 2014). Moreover, it was shown that loss of any 3M complex protein significantly altered the interphase microtubule network (Yan *et al.*, 2014). The core 3M-protein complex interacts with CUL9, which has been proposed to mediate the functions of the 3M complex via the ubiquitylation and degradation of survivin (Li *et al.*, 2014). The 3M-complex also interacts with the F box protein FBXW8, ROC1 and the tumour suppressor p53 (Yan *et al.*, 2014) suggesting it may contribute to correct cellular physiology through multiple mechanisms. In addition, mutations in *OBSL1*, *CUL7* and *CCDC8* are associated with altered IGFBP3 and IGF-1 signalling pathways suggesting the roles of 3M-complex in cellular growth (Hanson *et al.*, 2011). Interestingly, CCDC8 expression level was not significantly affected in wild type fibroblasts and those affected by 3M-syndrome with CUL7 and OBSL1 mutations suggesting that *OBSL1* and *CUL7* do not control the expression of *CCDC8* indicating its crucial role in cellular growth (Hanson *et al.*, 2011). In addition, *CCDC8* is known to uniquely regulate p53 mediated apoptotic process (Murray *et al.*). p53 is regulated by Tip60 acetylation without affecting the expression of p21 that is required for p53 mediated growth arrest (Dai *et al.*, 2011). p21 itself is not affected by CCDC8 knockdown; however, its knockdown represses the effect of PUMA, a P53 regulator of upregulated modulator of

apoptosis. Moreover, CCDC8 stimulates Lys120 acetylation of p53 by interacting with Tip60, suggesting the role of CCDC8 as an upstream regulator on interaction of Tip60-P53 (Dai *et al.*, 2011). This further underlines that *CCDC8* is required to regulate apoptosis mediated by *p53*. Taken together, *CCDC8*, an important component of 3M-complex, regulates microtubule dynamics and genome stability, cellular growth and growth signalling pathways as well as P53 mediated apoptosis.

The findings presented here indicate that epigenetic dysregulation of *BNC1* or *CCDC8* in breast tumours may contribute to metastasis to the brain and possibly other distant organs. *CCDC8* dysregulation occurs early during tumour evolution, in addition to being a potential therapeutic target this early inactivation has the potential to be utilised as a prognostic biomarker. Further analysis will be required including studies with larger sample numbers and studies to determine if such epigenetic markers can be discerned via non-invasive means such as analysis of circulating tumour material in the patients blood. *BNC1* promoter methylation and associated silencing is common in BBM but does not occur frequently in the originating breast tumours suggesting that their dysregulation may not necessarily benefit the primary tumour but are required for successful colonization of the brain. Further studies will be required to determine if these changes are detectable in circulating tumour cells, micrometastases, in other reservoir sites or only in macroscopic brain metastases. Our current understanding of the cellular function of these genes is far from complete. However, what is known about these two genes suggests that their dysregulation may be prognostic biomarker for BBM and are functionally significant, which could possibly represent novel therapeutic targets.

We investigated the association between the methylation status of *BNC1* and *CCDC8* with ER/PR/HER2 status of the primary tumours, which have metastasised to the brain in individual patients. *BNC1* was methylated in 6/8 BBM in matched pairs of samples (patient, 2, 3, 8, 12, 14 and 15) whereas it was methylated in only one of these associated primary tumours in individual patients *i.e.* patient 12 (Table 4.6). Some of these primary breast tumour patients where *BNC1* is not methylated are ER+/PR+/HER2- (such as patient 2, 3, and 12), ER+/PR-/HER2- (patient 10) and triple negative *i.e.* ER-/PR-/HER2- (patient 14). In this very small group of samples no association of BNC1 methylation with the receptor status of the patients was observed. Before any conclusions are made regarding this, further analysis with larger number of patients will be required. In addition, lymph node or vascular invasion and the stage of primary tumours do not follow any specific patterns of *BNC1* methylation. Interestingly, the two BBM tumours (patient 5 and 11) where *BNC1* is not methylated, took a relatively longer time to metastasise (*i.e.* Brain metastases surgery was taken place after 10 and 6 years from the removal of primary breast tumour respectively). This could possibly imply that the *BNC1* methylation may give metastasised tumours a growth advantage in the brain microenvironment. Methylation of *BNC1* in 6/8 BBM samples and its methylation in only one of the primary tumours in individual patients supports the idea that *BNC1* is not dysregulated in primary tumours to contribute to the process of metastatic evolution. However, its dysregulation due its promoter methylation may be crucial for survival of metastasised tumours in the brain (late event). Similarly, *CCDC8* was methylated in 10/11 BBM patients and all of their corresponding primary breast tumours. *CCDC8* methylation has occurred in the primary breast cancer patients regardless of their receptor status, tumour grade and lymph node or vascular invasion (Table 4.6). This supports our findings that the *CCDC8* methylation is an early event in

the process of BBM regardless of the clinical characteristics of the metastasising tumours.

Table 4.6: Clinical information of originating primary breast tumours and methylation status *BNC1* and *CCDC8*.

Patient	ER status	PR status	HER2 status	Grade and type	Lymph/Vascular Invasion	Duration between primary and BBM surgery	*BNC1*		*CCDC8*	
							Meth in PB	Meth in BBM	Meth in PB	Meth in BBM
Patient 1	Positive	Negative	Negative	Grade I, Invasive Ductal adenocarcinoma,	Vascular invasion noted	5 years				
Patient2	Positive (5-10%),	Positive (5-10%),	Negative	Infiltrating adenocarcinoma	NA	10 years				
Patient 3	Positive	Positive	Negative,	Grade III, IDC, metastatic,	Lymphovascular invasion	2 years				
Patient 5	Positive	Negative	Negative		NA	10 years				
Patient 8	NA	NA	NA	Grade II, IDC, Advanced metastatic carcinoma,	NA	2 years				
Patient 10	Negative	Negative	NA	Grade III IDC, infiltrating non-small cell carcinoma,	NA	3 years				
Patient 11	NA	NA	1+ (negative)	Grade III, infiltrating ductal carcinoma	NA	6 years				
Patient 12	Positive	Positive	1+ (negative)	Grade II, infiltrating ductal carcinoma	Lymphovascular invasion	5 years				
Patient 13	Negative	Negative	Negative	Grade III, infiltrating ductal carcinoma	NA	3 years				
Patient 14	Negative	Negative	Negative	Grade III, breast duct origin	NA	4 years				

Table 4.6: Clinical information of the primary breast tumours which metastasised to the brain. *BNC1* (methylated in 6/8 BBM) and *CCDC8* (methylated in 10/11 BBM) are frequently methylated in BBM patients. However, regardless of the clinical characteristics, BNC1 is methylated in only one of the corresponding primary tumours whereas *CCDC8* is commonly methylated in all the primary tumours in individual patients. This suggests the methylation status of *BNC1* and *CCDC8* may be independent of the clinical characteristics of the primary tumours. However, analyses of methylation in more patients are necessary to reach to this conclusion.

4.4 Conclusion

This analysis has identified two candidate metastatic suppressor genes (*BNC1* and *CCDC8*), which were frequently methylated and silenced in BBM, and infrequently methylated in unrelated primary breast tumours. These genes were identified from a screen of 85 candidate genes generated from a literature review. *CCDC8* was commonly methylated in BBM and their associated primary tumours in individual patients whereas *BNC1* was methylated in BBM and was not methylated in their associated primary tumours in individual patients. This suggests that dysregulation of *CCDC8,* that occurs in primary tumours (early events), could be a driver for metastasis of breast tumour cells to the brain whereas of *BNC1* dysregulation occurs after the tumour cells metastasised to the brain (late event) that contribute to survival of tumour cells in the brain microenvironment. The data suggests that *CCDC8* (a regulator of microtubule dynamics) and *BNC1* (a transcription factor with various target genes) could be used as prognostic markers or as therapeutic targets for BBM.

CHAPTER 5

Identifying novel candidate genes dysregulated in BBM by analyses of Genome wide 450K methylation array data from TCGA

5.1 Introduction

A comprehensive literature review generated a list of candidates, which are associated with one of the tumour types that readily metastases to the brain *i.e.* RCC, melanoma, breast and lung tumours (section 2.2.1) The main aim of the literature review was to identify candidate metastatic suppressor genes, which are methylated in either of the tumour types readily metastasising to the brain but not methylated in primary breast tumours. Hence, these genes may be infrequently methylated in primary breast tumours but are frequently methylated in RCC, melanoma or lung tumours. However, It is likely that genes identified by a literature review will not identify all potential candidate genes.

Among the primary tumours that metastasise to the brain, lung and breast metastasise to the brain more frequently (Soffietti *et al.*, 2002). Moreover, lung tumours metastasise more readily than the primary breast tumours (Feld *et al.*, 1984; Zhang *et al.*, 2013). We hypothesised that the genetic dysregulation responsible for the brain metastases may be commonly frequent in primary lung tumours and infrequent in primary breast tumours. In this regard, we hypothesised that genes that are infrequently methylated in non-metastasising breast tumours and frequently methylated in primary lung tumours (that readily metastasis to the brain) (Eichler *et al.*, 2011) may be found to be commonly methylated in metastatic brain tumours that derive from both lung and breast tumours

(Pangeni *et al.*, 2015).

The epigenetic events that are commonly frequent in lung, renal and melanoma tumours and drive early tumour development may also contribute to brain metastasis. However, as these events are rare in primary breast tumours they may not contribute to primary breast tumour formation but may contribute to BBM. Therefore, screening for genes frequently methylated in primary lung tumours and frequently methylated in breast tumours from TCGA was carried out to identify novel candidate genes dysregulated in BBM.

TCGA is a public repository of data on DNA copy number arrays, DNA methylation, exome sequencing, expression arrays, microRNA sequencing, single nucleotide polymorphisms (SNPs) of multiple cancer types (Cancer Genome Atlas Research, 2013). Many hundreds of human tumours have been investigated by TCGA to uncover molecular abnormalities at the genomic, transcriptomic, proteomic and epigenetic level resulting in an integrated picture of similarities and differences among tumour lineages (Cancer Genome Atlas, 2012). The analysis of TCGA data has revealed that breast cancers show novel gene dysregulation that contribute to tumour heterogeneity and clinically observable plasticity within the tumour populations (Cancer Genome Atlas Research, 2013). In this regard, it is possible that the genes screened by TCGA data analyses would identify potential candidate genes that contribute to breast cancer metastases to the brain.

The Infinium Methylation 450K BeadChip array has been designed to include over 485K CpG dinucleotides distributed across the genome, covering over 96% of CpG islands and 99% of the RefSeq genes from the UCSC database (Bibikova *et al.*, 2011;

Dedeurwaerder *et al.*, 2011). According to the HumanMethylation 450K array design, each gene is divided into a gene body, the 5' UTR, the first exon, gene body and the 3' UTR regions. The promoter region is further divided into TSS200 and TSS1500 located 200 and 1500 base pairs upstream of the transcription start site (TSS) (Bibikova *et al.*, 2011). The candidate gene approach described in Chapter 4 has been limited to those candidate genes which are known to be associated with primary tumours. Therefore, the analysis of the TCGA data was carried out to find additional, novel candidate metastatic suppressor genes that may contribute to BBM.

Unlike the candidate gene approach which identified known candidate genes already associated with primary tumour types (section 4), this analysis was intended to identify novel candidate metastatic suppressor genes distributed across the genome. To ensure that promoter-associated CpG islands (see 1.2.1.1 for CpG islands details) were identified, only those probes that are located in the 5' region of the gene (5'UTR) or up to 1500 base pairs from the transcription start site (TSS, TSS200, TSS1500) were selected. The analysis was carried out from the methylation array data for each of 20 primary lung and breast tumours (with no evidence of metastasis) downloaded from the TCGA (see section 3.7.1, table 3.6). The methylation status of individual probes for each sample was compared between primary lung and breast tumours. Probes that are not methylated (β value ≤ 0.25) in primary breast tumours and methylated (β value ≥ 0.60) in primary lung tumours identified 4 candidates that were further validated using CoBRA.

5.2 Results:

5.2.1 Bioinformatic analyses of breast and lung tumours from TCGA

The objective of this study was to identify genes dysregulated due to promoter hypermethylation. Therefore the probes which are not associated with promoter region, were discarded. The selection of probes that are located in and around the promoter region of the gene (TSS-TSS1500), generated a list of 199590 probes. β value of 0.5 in a gene denotes that at least one of its allele is methylated. Therefore, to be statistically stringent, only those probes with β≥0.75 were considered as methylated (for lung tumours). Similarly, only those probes with β value ≤ 0.25 were considered as unmethylated (for breast tumours) to make sure that neither of the alleles of a gene is methylated. Therefore, screening of only those probes unmethylated (β value ≤ 0.25) at least in 15/20 (75%) of primary breast tumours generated a long list of 97155 probes. Similarly, screening of only probes methylated (β value ≥0.75) in at least 15/20 lung tumours generated a list of 24052 probes. Comparison of these two sets of probes *i.e.* commonly unmethylated in primary breast tumours and methylated in lung tumours generated no probes in common. Therefore, multiple rounds of analyses were carried out to generate three independent lists of probes for lung tumours by considering β value of ≥0.70, ≥0.65 and ≥0.60 in 15/20 (75%) of tumours that generated a long list of 28396, 32584 and 36652 probes respectively (Figure 5.1).

Multiple rounds of analyses were carried out to identify probes common in lung and breast tumours (Figure 5.1). Comparison of 97155 unmethylated probes in breast

tumours and 36652 methylated probes in lung tumours, (β value of ≥0.60) generated only eight probes in common that corresponded to six genes (*GALNT9*, *KRT222*, *PLEKHA6*, *TFAP2A*, *TSPAN4* and *ZNF808*) (Figure 5.1 and table 5.1). One of these genes *KRT222* did not have a well-defined CpG island (Gardiner-Garden, 1987). Therefore, this genome wide approach identified five candidate genes (*GALNT9*, *TFAP2A*, *TSPAN4, PLEKHA6* and *ZNF808*), which were analysed further using CoBRA in an initial cohort of 15 BBM samples.

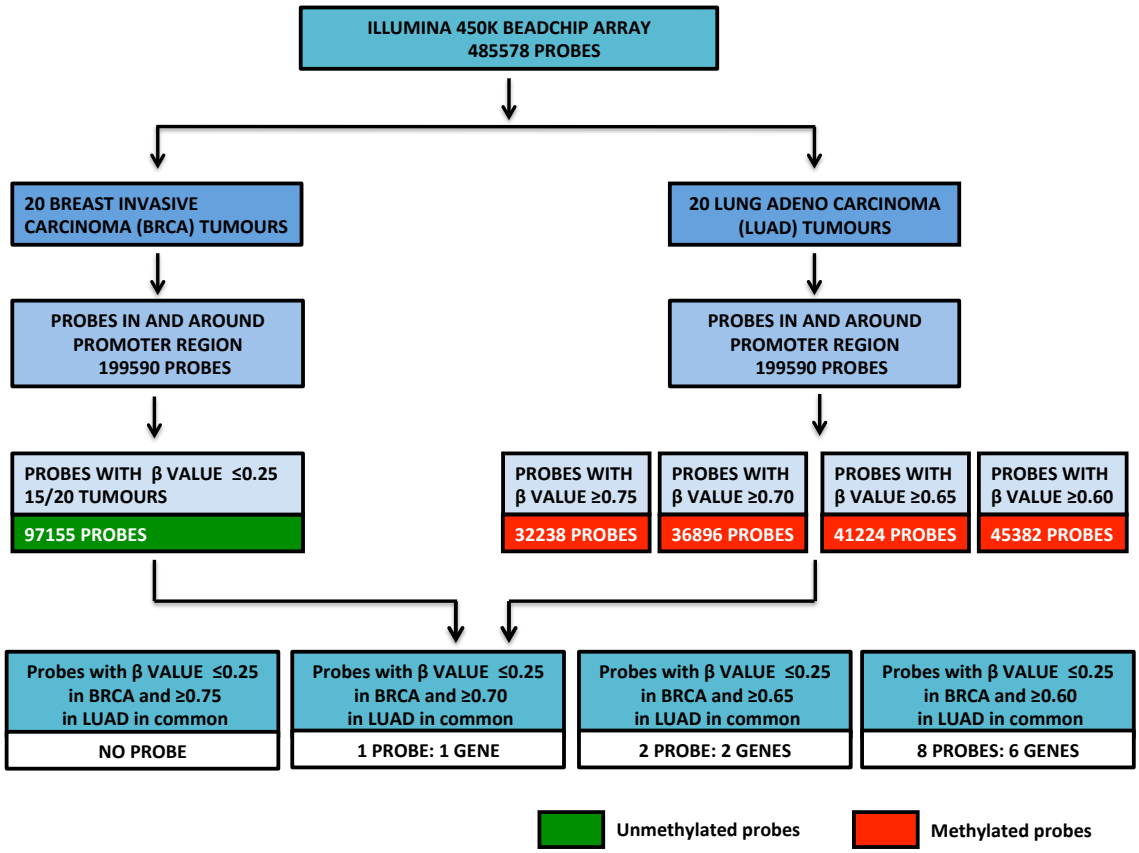


Figure 5.1: Bioinformatic analyses of TCGA data. Analyses of 450K methylation array data for 20 primary breast tumours (Breast Invasive Carcinoma; BRCA) and 20 primary tumours (Lung Adenocarcinoma; LUAD) were carried out. The analysis was carried out to screen probes/genes, which are unmethylated in primary breast tumours (β value ≤ 0.25 in 10/20 tumours) and methylated in primary lung tumours (β value ≥0.75 in 15/20 tumours) in common. The analyses identified 6 genes (8 probes), which were then analysed independently in the laboratory.

Table 5.1: Genes identified from TCGA data analyses that may contribute to breast to brain metastases (BBM)

BeadChip Probe	Gene symbol	Accession	Probe position relative to gene	Synonyms	Full name	Function
cg09879122 cg27483007	*GALNT9*	NM_001122636.1	TSS200	*GalNAc TRANSFERASE 9; GalNAcT9*	*N-acetyl galactosaminyl transferase 9*	Catalyzes O-glycosylation (Shinya *et al.*, 2000; Berois *et al.*, 2013)
cg11641791	*KRT222*	NM_152349.2	TSS200	*KA21; KRT222P*	*Keratin 222, Keratin like protein 222*	Structural component of cytoskeleton (Rappaport *et al.*, 2014)
cg25407979	*PLEKHA6*	NM_014935.4	5'UTR	*PEPP3; PEPP-3*	*pleckstrin homology domain containing, family A member 6*	Involved in cellular signalling and cytoskeleton organisation (Spellmann *et al.*, 2014)
cg10899301 cg24902920	*TFAP2A*	NM_003220.2	TSS1500	*AP-2; BOFS; AP2TF; TFAP2; AP-2alpha*	*TF activating enhancer binding protein 2 alpha*	Required for neural crest induction (Li & Cornell, 2007) Development and differential, cellular and pathological process, reduced function causes development diseases, abnormal expression in various human cancer,
cg14277925	*TSPAN4*	NM_001025237.1	5'UTR	*NAG2; NAG-2; TM4SF7; TSPAN-4; TETRASPAN*	*tetraspanin 4*	Cellular growth, adhesion and differentiation (Todd *et al.*, 1998a)
cg22455450	*ZNF808*	NM_001039886.3	5'UTR		*Zink Finger protein 808*	Zinc finger protein, may be involved in transcriptional regulation (Lancet *et al.*, 2013)

Table 5.1: Probes/Genes identified from bioinformatics analysis of TCGA data to screen novel genes that may contribute to BBM. The analyses identified eight probes (six genes) in common that are unmethylated in primary breast tumours and methylated in primary lung tumours. The genes identified are involved in various biological functions such as posttranslational modifications, cellular signaling, cytoskeleton, and transcriptional regulations and are associated with human cancers.

5.2.2 Methylation analyses of five BBM candidate genes

CoBRA was carried out to investigate the methylation status of five genes (*GALNT9*, *TFAP2A*, *TSPAN4, PLEKHA6* and *ZNF808*) in an initial cohort of 15 BBM samples. Only those genes, which were frequently methylated in an initial cohort of 15 BBM samples, were used for further analyses. *GALNT9* was frequently methylated (8/15 (55%)) in BBM samples, the other four genes *TFAP2A*, *TSPAN4, PLEKHA6* and *ZNF808* were not methylated in any of the BBM samples, and were excluded from further analysis.

The CoBRA primers used to amplify promoter region of these genes are listed in appendix A3.1.

Promoter methylation status of *GALNT9* was carried out in a second cohort of 15 BBM samples. Therefore, methylation status of *GALNT9* was determined in a total of 30 BBM samples. However, no amplified PCR product could be generated for two of the BBM. *GALNT9* was methylated in 16/29 (55%) of the BBM samples (figure 5.2). The observed high frequency of methylation merited further analysis of *GALNT9* promoter methylation status in a cohort of unrelated primary breast tumours. A quantitative methylation analysis of GALNT9 is given in section 5.2.6.

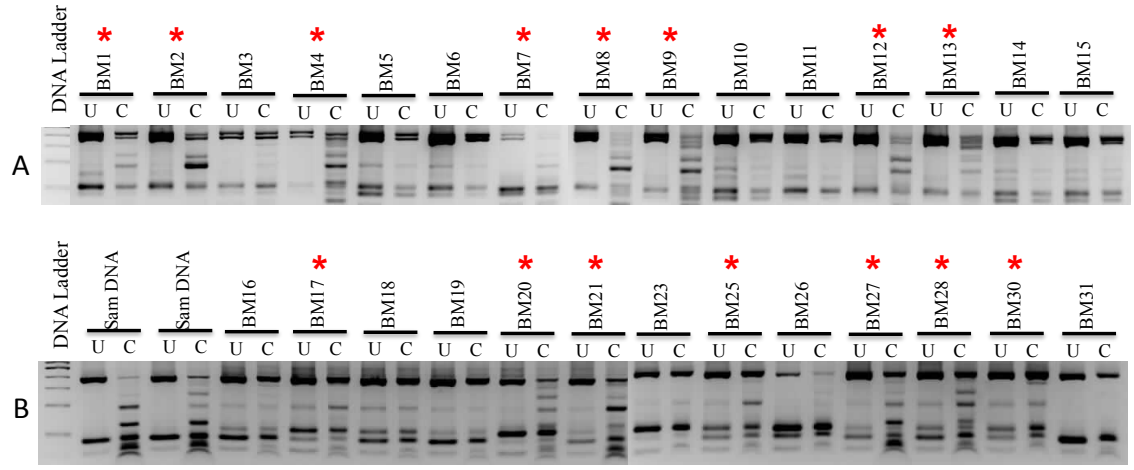


Figure 5.2: Methylation status of *GALNT9* in BBM. *GALNT9* is methylated in 17/29 (58%) of BBM samples. Promoter region methylation in of 15 samples (A and B) was carried out using CoBRA. Sam DNA: fully methylated positive control, BM: Brain metastases, U: Uncut/control sample, C: cut by restriction enzyme, *: methylated samples.

5.2.3 Methylation analyses of *GALNT9* in unrelated primary breast tumours

Methylation status of *GALNT9* in an independent cohort of 40 primary breast tumours with no evidence of distant metastases was carried out. The clinical information on these primary tumours is given in appendix B4. It was not possible to amplify *GALNT9* in seven of the primary samples. *GALNT9* (which was methylated in 58% of BBM) was not methylated in any of the 33 primary breast tumours analysed (Figure 5.3). Therefore, there was a significant difference in promoter methylation status of *GALNT9* between unrelated primary breast tumours and BBM ($p=0.0001$).

Table 5.2: Methylation status of genes identified from TCGA data analyses in breast to brain metastases (BBM)

Gene	Breast to Brain metastases (BBM) samples (n=15)															Meth	% Meth (n=15)
	BM1	BM2	BM3	BM4	BM5	BM6	BM7	BM8	BM9	BM10	BM11	BM12	BM13	BM14	BM15		
GALTN9																8	53
PLEKHA																0	0
TFAP2A																0	0
TSPAN4																0	0
ZNF808																0	0

Methylation Status of *GALNT9* in second cohort of BBM samples (n=15)																	
Gene	BM16	BM17	BM18	BM19	BM20	BM21	BM22	BM23	BM24	BM25	BM26	BM27	BM28	BM29	BM30	BM31	Meth	% meth (n=29)
GALNT9																	9	58

Table 5.2: Promoter methylation status of five candidate genes identified from TCGA data analyse. Initially, a cohort of 15 BBM samples was used to investigate the methylation status of the genes. Only *GALNT9* was frequently methylated in these BBM samples, following analysis of a further 15 samples the promoter methylation frequency of *GALNT9* was determined to be 58% (17 of 29 samples).

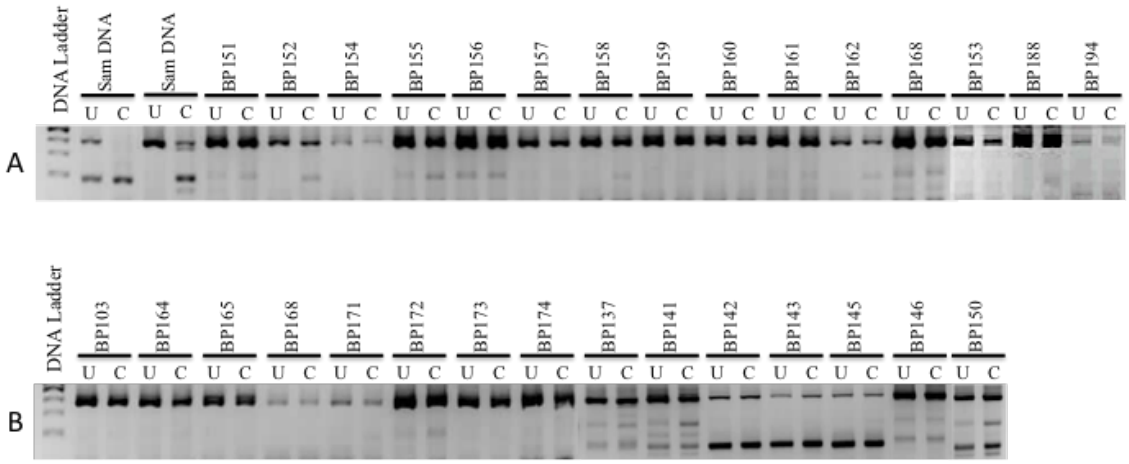


Figure 5.3: Methylation status of *GALNT9* in primary breast tumours. *GALNT9* is NOT methylated in any of the primary tumours. Promoter region methylation in 40 unrelated primary breast tumours was carried out using CoBRA. Sam DNA: fully methylated positive control, BP: primary breast tumours, U: Uncut/control sample, C: cut by restriction enzyme, *: methylated samples.

5.2.4 Expression analysis of *GALNT9* in metastatic brain tumours

As *GALNT9* was frequently methylated in BBM the next task was to investigate if promoter methylation of *GALNT9* was associated with its gene silencing. Only 15 BBM samples were available for RNA extraction among 30 breast to brain metastases. Total RNA was extracted from these samples to determine the expression of *GALNT9* by RT-PCR (Figure 5.4). RT-PCR shows that the expression of *GALNT9* correlates with its promoter methylation in BBM. *GALNT9* is frequently silenced (**) in 5 methylated tumours (BM12, BM20, BM27, BM28 and BM29), downregulated (*) in 3 methylated tumours (BM13, BM17 and BM30), and is expressed in 4 unmethylated tumours (BM11, BM14, BM18 and BM23). However, *GALNT9* is silenced in three BBM samples in which its promoter is unmethylated (BM3, BM15 and BM16). The silencing of *GALNT9* in unmethylated samples could be due to the genetic alteration other than the DNA methylation such as deletion. (See section 5.4 for quantitative methylation and expression analyses).

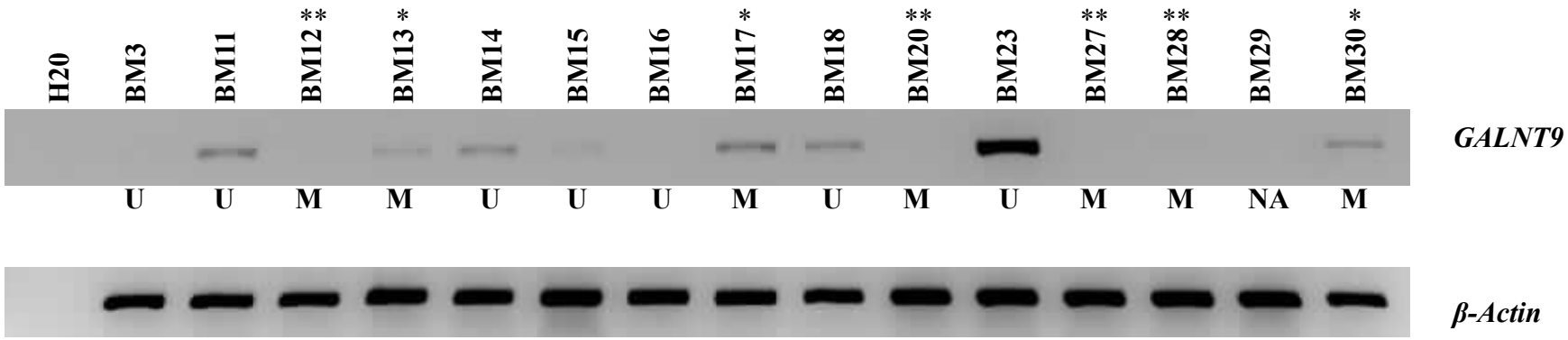


Figure 5.4: Gel electrophoresis of reverse transcription PCR (RT-PCR) of GALNT9 in BBM. Expression analysis using RT-PCR shows that the expression of *GALNT9* correlates with its promoter methylation in BBM. *GALNT9* is expressed in unmethylated (U), downregulated (*) in partially methylated, and is silenced (**) in fully methylated tumours. (See figure 5.2 for methylation analysis). Expression of β-actin was determined to ensure equal loading for all samples. Of 15 BBM samples available for expression analysis, four unmethylated samples (BM11, BM14, BM18 and BM23) are expressed; three partially methylated samples (BM13, BM17 and BM30) are downregulated and five methylated samples (BM12, BM20, BM27, BM28 and BM29) are silenced. Three unmethylated samples (BM3, BM15 and BM16) are also silenced which could be due to other genetic aberrations such as deletion or dysregulation on complex network of the gene on those samples. (See section 5.4 for quantitative expression analysis)

5.2.5 Promoter methylation status of *GALNT9* in brain metastases and associated primary breast tumours from individual patients

We analysed the methylation status of *GALNT9* in matched pairs of tumours *i.e.* metastatic brain tumours and corresponding primary tumours from individual patients. We have 10 of these pairs, however, some loci in the primary tumour DNA proved refractive to amplification. It was possible to amplify the *GALNT9* promoter region in 6 of these samples. Methylation was observed in 0/6 of these primary breast tumour samples and in 4/6 of the resulting brain metastases (Figure 5.5). These results suggest that *GALNT9* promoter methylation occurs at a late stage in the evolution of metastatic brain tumours, possibly after the primary tumour cells have metastasised to the brain. Alternatively, methylation of these genes may occur in a small subset of cells within the primary tumour (below the detection threshold of this assay) and these cells are enriched in the metastatic tumour.

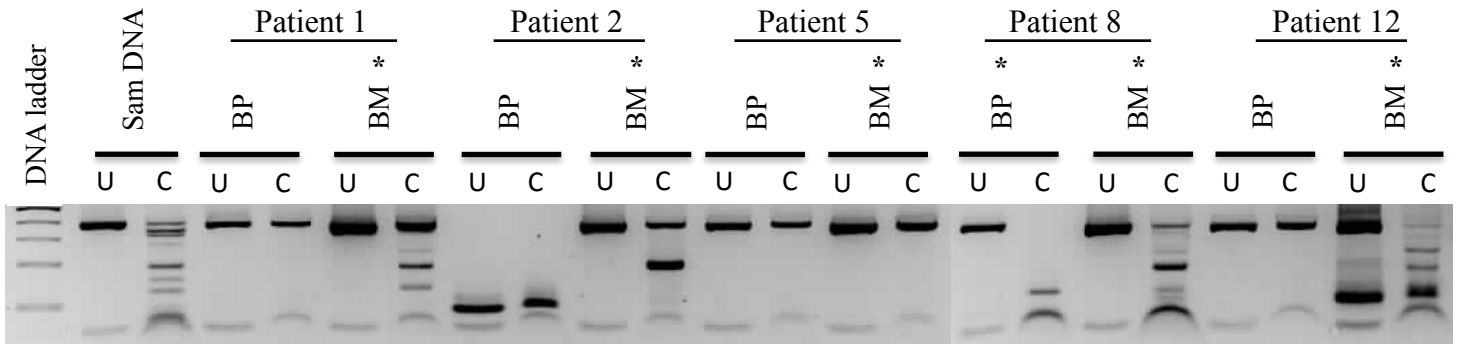


Figure 5.5: Methylation status of *GALNT9* in metastatic brain tumours and their corresponding originating primary breast tumours from individual patients. CoBRA was used to determine the methylation status; small, digested PCR products in the BstuI cut (C) lane compared to the undigested (U) lane indicates promoter methylation in a sample. *GALNT9* was frequently methylated (*) in metastatic brain tumours (BM) and are NOT methylated in the originating breast primary (BP) tumours. Of 5 matched pairs analysed, *GALNT9* is methylated in 4 metastatic brain tumours (see patient 1, 2, 8 and 12), whereas it is methylated in only one of the corresponding primary tumour (patient 8). Sam DNA: fully methylated positive control, BP: Breast Primary tumour, BM: Metastatic Brain tumour, BN: adjacent Normal Breast tissue, U: Uncut/Control sample, C: cut by methylation specific restriction enzyme, *: Methylated sample

5.2.6 Quantitative methylation analysis for *BNC1*, *CCDC8* and *GALNT9*

Quantitative methylation analyses for *BNC1*, *CCDC8* and *GALNT9* were carried out to ensure that that CoBRA digests were representative of high methylation status in tumours. For this purpose, the promoter region amplified by CoBRA was cloned (see section 3.3.5 for detailed methodology) and sequenced to carry out base-resolution analysis of promoter region methylation for these genes. The sequencing was used to determine the bisulphite-modified alleles in each individual tumour. Based on the methylation status of individual CpG dinucleotides in each allele, the methylation index (MI) of CpG islands (see figure 5.6 for the CpG island region amplified) for individual tumours was determined. MI is defined as the total number of methylated CpG dinucleotides given as a percentage of all CpGs analysed. The MI for promoter regions, which were determined by CoBRA as methylated ranged from 60 to 91 % whereas those promoters deemed not to be methylated by CoBRA, had MIs ranging between 0 and 36 %. From this analysis, we have defined that, for these samples, physiologically significant methylation levels are those of ≥60 % MI and lack of physiologically significant methylation is defined as <40 % MI (figure 5.7). The detailed sequencing of alleles in individual tumours, are given in appendix C1.

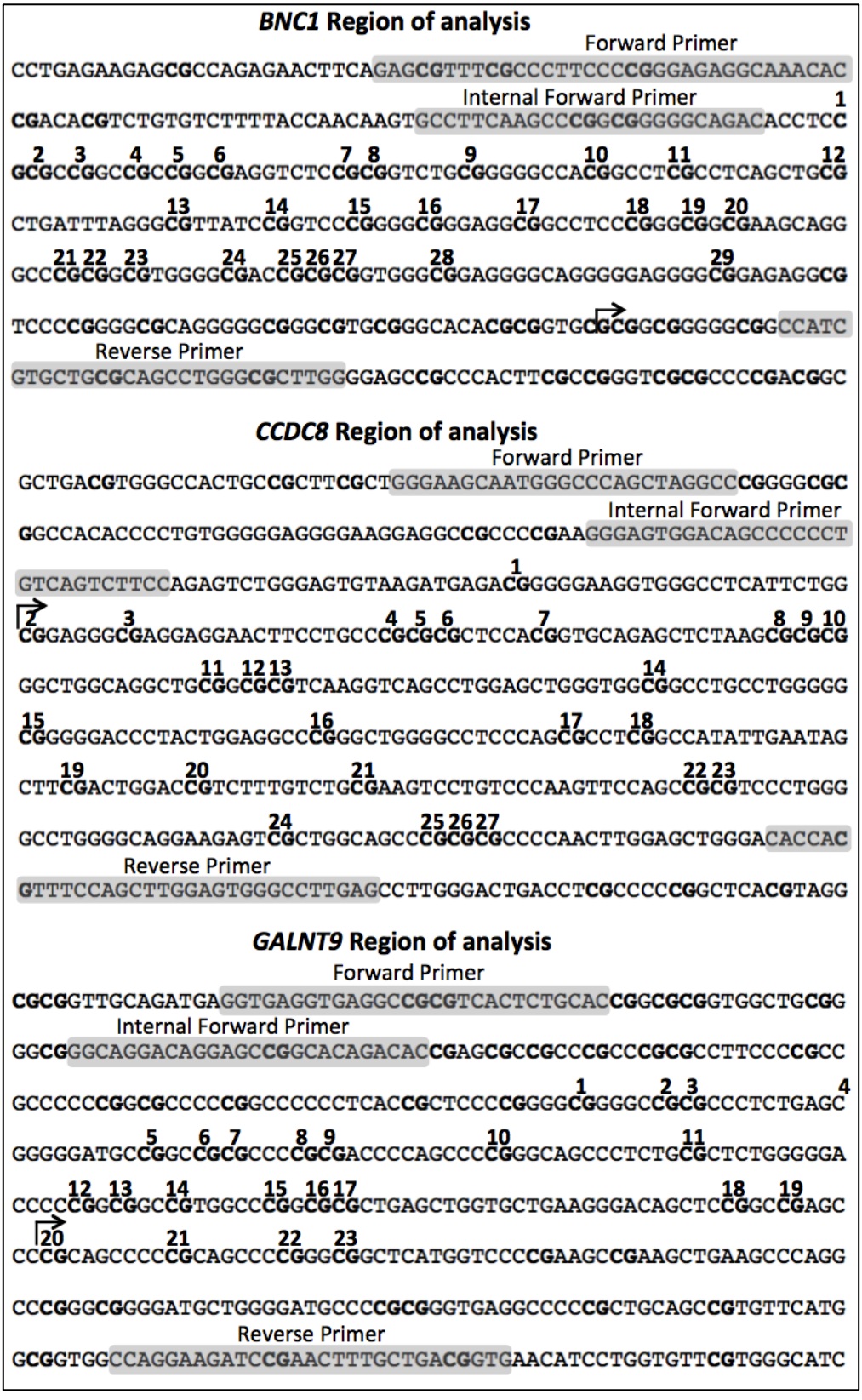

BNC1 Region of analysis
Forward Primer
CCTGAGAAGAGCGCCAGAGAACTTCAGAGCGTTTCGCCCTTCCCCGGGAGAGGCAAACAC
Internal Forward Primer
1
CGACACGTCTGTGTCTTTTACCAACAAGTGCCTTCAAGCCCGGCGGGGGCAGACACCTCC
2 3 4 5 6 7 8 9 10 11 12
GCGCCGGCCGCCGGCGAGGTCTCCGCGGTCTGCGGGGGCCACGGCCTCGCCTCAGCTGCG
13 14 15 16 17 18 19 20
CTGATTTAGGGCGTTATCCGGTCCCGGGGCGGGAGGCGGCCTCCCGGGCGGCGAAGCAGG
21 22 23 24 25 26 27 28 29
GCCCGCGGCGTGGGGCGACCGCGCGGTGGGCGGAGGGGCAGGGGGAGGGGCGGAGAGGCG
TCCCCGGGGCGCAGGGGGCGGGCGTGCGGGCACACGCGGTGCGCGGCGGGGGCGGCCATC
Reverse Primer
GTGCTGCGCAGCCTGGGCGCTTGGGGAGCCGCCCACTTCGCCGGGTCGCGCCCCGACGGC
CCDC8 Region of analysis
Forward Primer
GCTGACGTGGGCCACTGCCGCTTCGCTGGGAAGCAATGGGCCCAGCTAGGCCCGGGGCGC
Internal Forward Primer
GGCCACACCCCTGTGGGGGAGGGGAAGGAGGCCGCCCCGAAGGGAGTGGACAGCCCCCCT
1
GTCAGTCTTCCAGAGTCTGGGAGTGTAAGATGAGACGGGGGAAGGTGGGCCTCATTCTGG
2 3 4 5 6 7 8 9 10
CGGAGGGCGAGGAGGAACTTCCTGCCCGCGCGCTCCACGGTGCAGAGCTCTAAGCGCGCG
11 12 13 14
GGCTGGCAGGCTGCGGCGCGTCAAGGTCAGCCTGGAGCTGGGTGGCGGCCTGCCTGGGGG
15 16 17 18
CGGGGGACCCTACTGGAGGCCCGGGCTGGGGCCTCCCAGCGCCTCGGCCATATTGAATAG
19 20 21 22 23
CTTCGACTGGACCGTCTTTGTCTGCGAAGTCCTGTCCCAAGTTCCAGCCGCGTCCCTGGG
24 25 26 27
GCCTGGGGCAGGAAGAGTCGCTGGCAGCCCGCGCGCCCCAACTTGGAGCTGGGACACCAC
Reverse Primer
GTTTCCAGCTTGGAGTGGGCCTTGAGCCTTGGGACTGACCTCGCCCCCGGCTCACGTAGG
GALNT9 Region of analysis
Forward Primer
CGCGGTTGCAGATGAGGTGAGGTGAGGCCGCGTCACTCTGCACCGGCGCGGTGGCTGCGG
Internal Forward Primer
GGCGGGCAGGACAGGAGCCGGCACAGACACCGAGCGCCGCCCGCCCGCGCCTTCCCCGCC
1 2 3 4
GCCCCCCGGCGCCCCCGGCCCCCCTCACCGCTCCCCGGGGCGGGGCCGCGCCCTCTGAGC
5 6 7 8 9 10 11
GGGGGATGCCGGCCGCGCCCCGCGACCCCAGCCCCGGGCAGCCCTCTGCGCTCTGGGGGA
12 13 14 15 16 17 18 19
CCCCCGGCGGGCCGTGGCCCGGCGCGCTGAGCTGGTGCTGAAGGGACAGCTCCGGCCGAGC
20 21 22 23
CCCGCAGCCCCCGCAGCCCCGGGCGGCTCATGGTCCCCGAAGCCGAAGCTGAAGCCCAGG
CCCGGGCGGGGATGCTGGGGATGCCCCGCGGGTGAGGCCCCCGCTGCAGCCGTGTTCATG
Reverse Primer
GCGGTGGCCAGGAAGATCCGAACTTTGCTGACGGTGAACATCCTGGTGTTCGTGGGCATC

Figure 5.6: The promoter region/CpG islands of *BNC1*, *CCDC8* and *GALNT9*. The region amplified for CoBRA analysis is found between the Internal Forward primer and the Reverse primer. CpG dinucleotides are highlighted in bold. CpG dinucleotides analysed by cloning and sequencing of individual alleles are numbered. An arrow indicates the transcription start site.

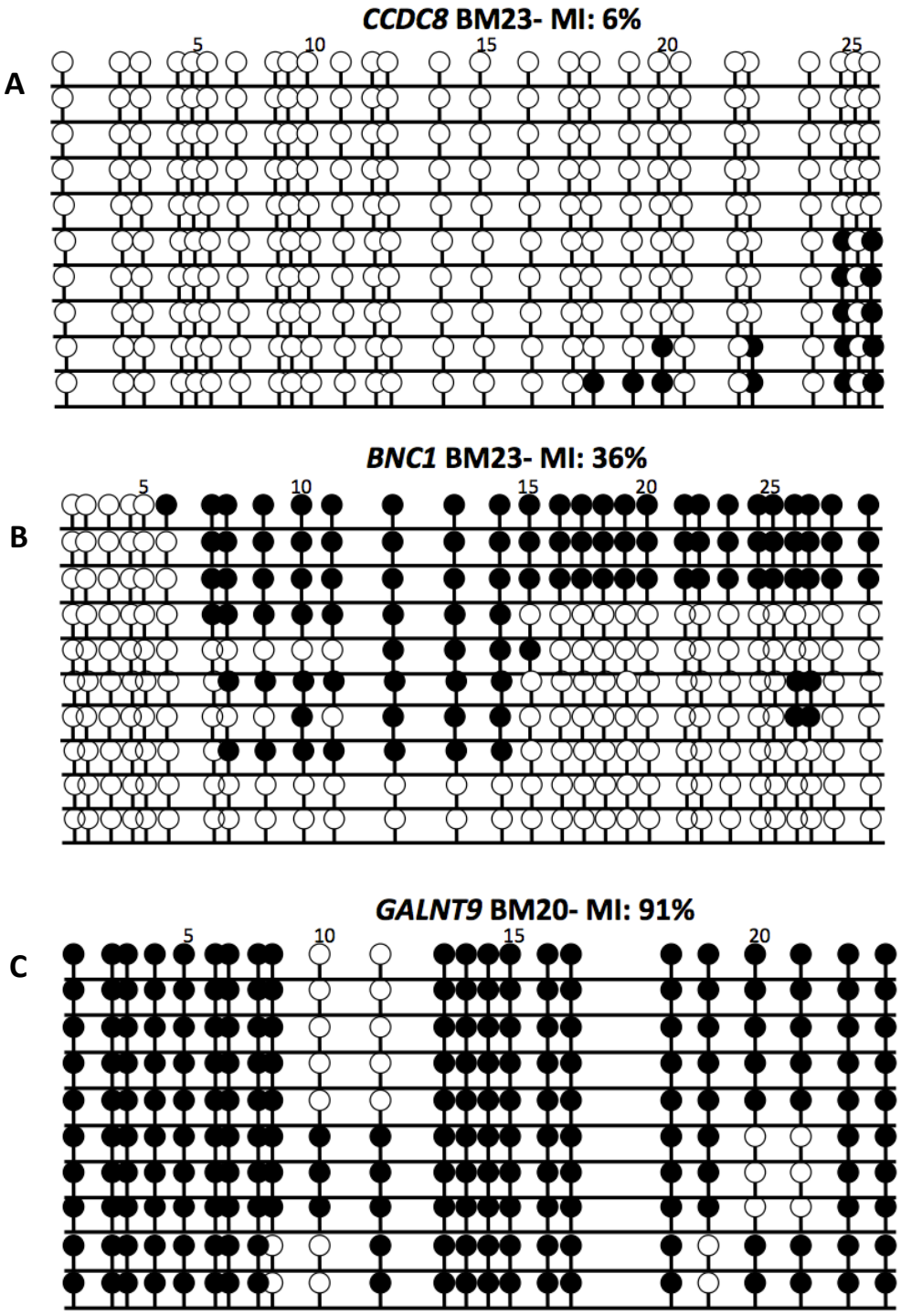


Figure 5.7: Bisulphite sequencing of individual alleles from tumours. Tumours were analysed by cloning and sequencing bisulphite-PCR products to determine the extent of methylation within the region analysed by CoBRA. 10 clones/alleles were sequenced for each tumour and the methylation index (MI) for each tumour determined. (A) A representative example of tumours that were determined to

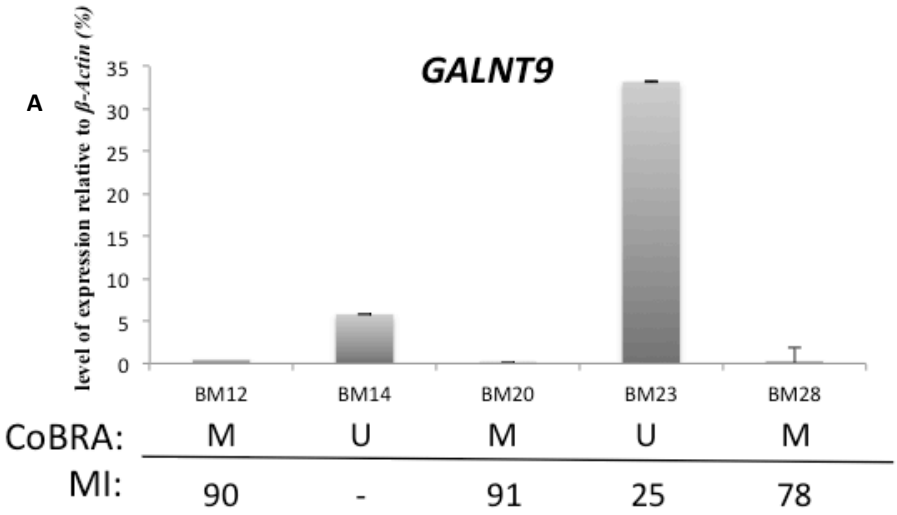
have significant *CCDC8* promoter methylation by CoBRA having methylation index of only 6%. Similarly, the same tumour has a methylation index of 36% for *BNC1* (B), and a representative example of tumours that has methylation index of 91% for *GALNT9* (C). MI is defined as the total number of methylated CpG dinucleotides given as a percentage of all CpGs analysed. The details sequencing of the region amplified for these genes for individual tumour is given in C1.

5.2.7 The Quantitative Expression analyses for *BNC1*, *CCDC8* and *GALNT9*

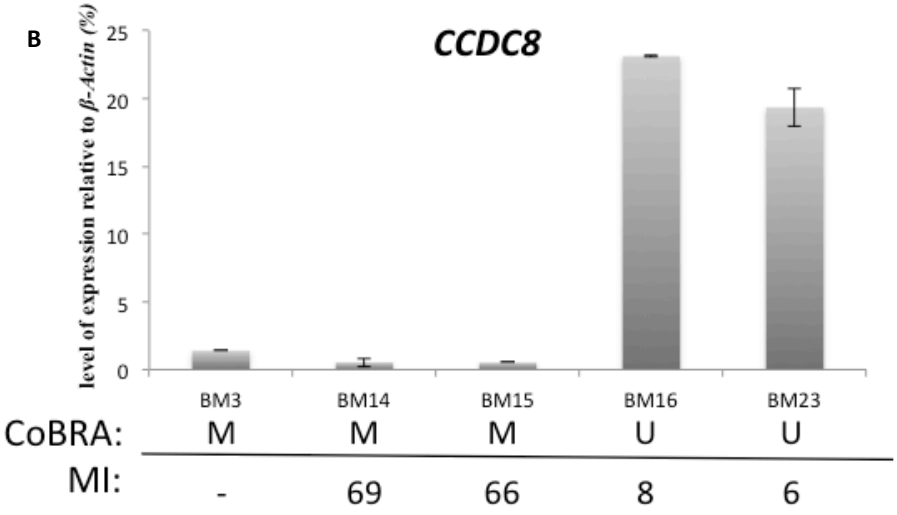
The expression level of each gene was quantified in relation to the expression of β-actin, in tumours with unmethylated promoters (MI = 0–25 %). The maximum expression of these genes was 49, 23 and 33 % that of β-actin, respectively. *BNC1*, *CCDC8* and *GALNT9* were frequently downregulated or silenced in these tumours and reduced expression correlated to promoter methylation as determined by CoBRA and base-resolution sequencing (see figure 5.8 and appendix C2). These genes were also commonly silenced in breast cancer cell lines; this silencing was reversed following treatment with 5-Aza- 2′-deoxycytidine an inhibitor of DNA methyltransferase enzymes (Patra & Bettuzzi, 2009).

This common *CCDC8* methylation in primary breast tumour and resulting brain metastasis was confirmed by sequencing individual alleles for pairs of tumours from two patients, patient 11 and 15 (BM11, Primary BP 11 and BM15, Primary BP 15). Both primary tumour DNA and BM DNA were found to have MIs above 73 %

A


GALNT9
level of expression relative to β-Actin (%)
35
30
25
20
15
10
5
0
BM12
BM14
BM20
BM23
BM28
CoBRA:
M
U
M
U
M
MI:
90
-
91
25
78

B


CCDC8
level of expression relative to β-Actin (%)
25
20
15
10
5
0
BM3
BM14
BM15
BM16
BM23
CoBRA:
M
M
M
U
U
MI:
-
69
66
8
6

C

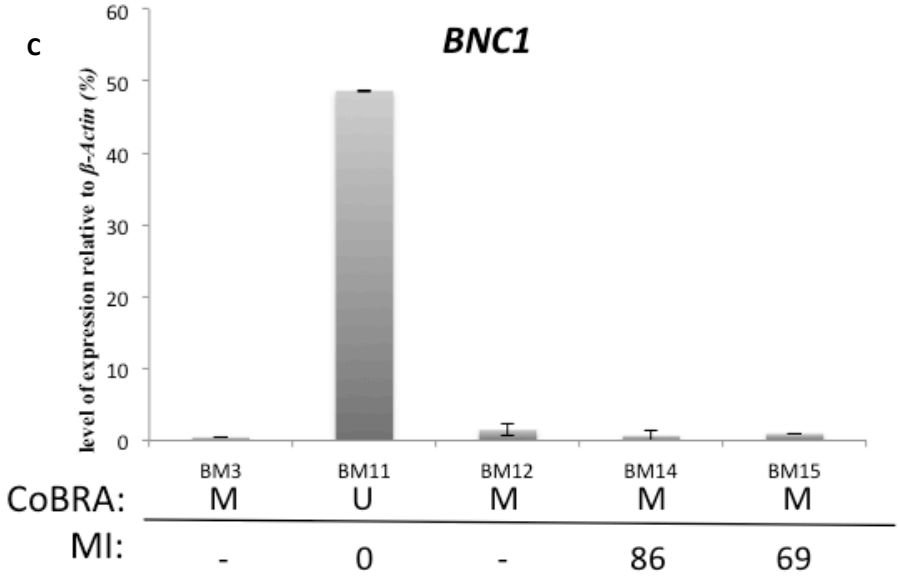

BNC1
level of expression relative to β-Actin (%)
60
50
40
30
20
10
0
BM3
BM11
BM12
BM14
BM15
CoBRA:
M
U
M
M
M
MI:
-
0
-
86
69

Figure 5.8: Expression levels of each gene were quantified in relation to the expression of β-actin. The methylation status was determined by either CoBRA or sequencing of individual alleles to determine the methylation index (MI) for individual tumours. High levels of expression were not associated with high levels of methylation in the region analysed. (BM brain metastasis, M methylated, U unmethylated, '-'analysis was not done). For full set of expression, see appendix C2.

5.2.8 Methylation status of *BNC1*, *CCDC8* and *GALNT9* in metastatic and non-metastatic primary breast tumours from the TCGA

We downloaded all the clinical information available for the metastatic primary breast tumours from TCGA. Clinical information from TCGA (table 5.3) revealed that there were only 21 primary breast tumours (BRCA), which were predicted to be metastatic to distant sites such as lung, breast, liver and skin. However, methylation data was available for only 14 of these metastatic tumours (Table 5.4). The methylation status of probes that are located within the promoter region amplified for each gene (*BNC1*, *CCDC8* and *GALNT9*) were investigated in the clinical samples from TCGA. There were only two probes for *BNC1* that were within the region amplified (Figure 5.9A). One of these probes (cg18560204) was methylated ((β value ≥ 0.70) in only one of 14 metastatic tumours *i.e.* tumour M12 (table 5.5) where as the second probe (cg27304406) was not methylated in any of the metastatic tumours. Similarly, there are two probes for *GALNT9* within a promoter region amplified (figure 5.9C); these probes are not methylated in primary tumours (with an exception of a probe, cg12075445 methylated in a single tumour). Therefore, this supports our findings that *GALNT9* methylation occurs late only in brain metastases. Furthermore, there were two probes for *CCDC8* located within the promoter region (Figure 5.9B); however, the methylation data was available only for a probe (cg06747432). This probe cg06747432 for *CCDC8* is methylated in 4/14 tumours; M1, M8, M11 and M12. M1 and M8 are stage III and stage II tumours with unknown distance metastases respectively. Methylation of *CCDC8* in these primary tumours may be an early event for general metastasis, or possibly brain specific metastases (metastases in the brain might not have been evaluated or detected). Moreover, M11 and M12 are stage IV and stage III tumours with liver and bone

metastases respectively. These tumours initially might have metastasised to the liver/bone and have evolved as a detectable macrometastases at the time of surgery. There could possibly be undetectable micrometastases in the brain. It is important to note that there might have been different genomic alterations in these primary breast tumours (M11 and M12) responsible for liver and bone metastases respectively (not *CCDC8* methylation) in a small subset of tumour cells due to the heterogeneous nature of a tumour and *CCDC8* methylation might have occurred in another subset of tumour cells with an advantage (such as gaining metastatic potential due to *CCDC8* methylation) to invade into the brain. These data supports our finding that *CCDC8* methylation occurs in primary breast tumours that eventually evolve into brain metastases.

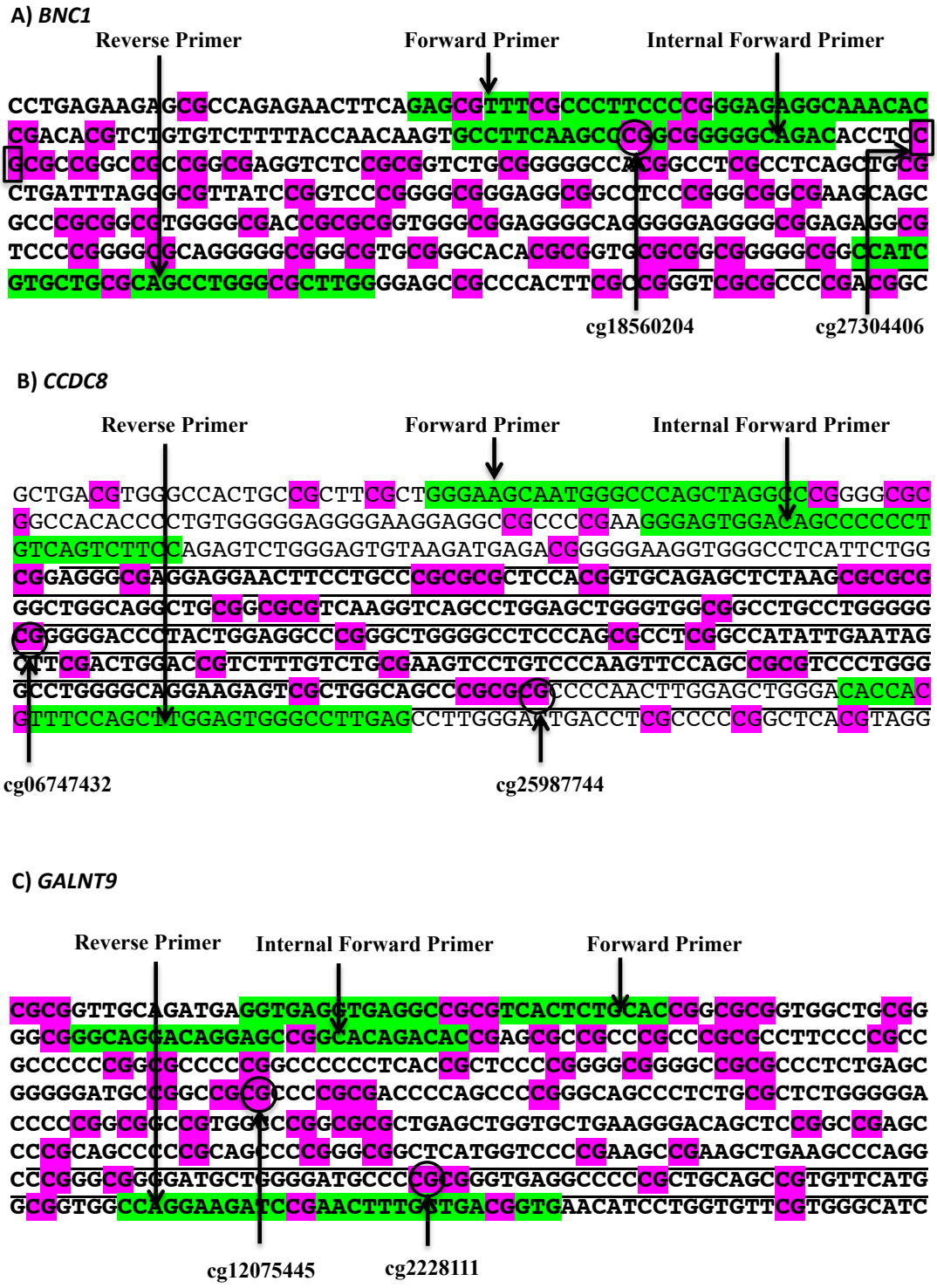


Figure 5.9: CoBRA primers designed for *BNC1*, *CCDC8* and *GALNT9* to investigate the methylation status of promoter region of each gene. The probes from 450K methylation array that are located within the region amplified in these genes have been identified to determine the methylation status of individual probes.

The central CG for each probe is marked in a circle (purple). Other independent CpGs (CGs) are shaded purple. The primers are shaded in green. The transcripts are underlined.

Table 5.3: Clinical information of metastatic primary breast tumours from TCGA.

Tumour Code	Patient's age	ER STATUS	PR STATUS	HER2 Status	micrometastasis	distant metastasis	Distant metastasis sites	Tumour stage	Pathological stage (metastasis)	Pathological stage (Lymph node)
M1	62	Negative	Negative	NA	YES	YES	NA	Stage IIIA	M0	N2a
M2	58	Negative	Negative	NA	NO	YES	Lung/Bone/Liver	Stage IIIC	M0	N1mi
M3	68	Negative	Negative	2+	NO	YES	Lung	Stage IIIC	M0	N0 (i-)
M4	55	Negative	Negative	2+	NA	YES	Bon/Liver	Stage III	M0	N1b
M5	69	Positive	Positive	NA	YES	YES	Bone	Stage I	M0	N3
M6	46	Positive	Positive	2+	NO	YES	Bone	Stage IIB	MX	N3b
M7	87	NA	NA	NA	YES	YES	Other	Stage IIB	M0	N3b
M8	55	Positive	Negative	NA	NO	YES	NA	Stage IIB	MX	N1b
M9	47	Positive	Negative	NA	NA	YES	Bone	NA	NA	NA
M10	82	Positive	Negative	NA	NO	NA	Liver	Stage IV	M1	N2a
M11	66	Negative	Negative	0	NO	NA	Liver	Stage IV	M1	N2
M12	43	Positive	Positive	2+	NA	NA	Bone	Stage IIIA	M1	N3
M13	47	Positive	Positive	1+	NO	NA	Bone/Liver	Stage IIIA	M0	N2
M14	63	Positive	Positive	2+	NO	NA	Bone	Stage IV	M0	N2

Table 5.3: Clinical information of metastatic BRCA samples from TCGA has shown that the distant metastasis of primary breast tumours (in the data sets available so far) has taken place only in the lungs, bone and liver. No metastasis to the brain has been specified. The methylation data for *BNC1*, *CCDC8* and *GALNT9* on these tumours is given in Table 5.6. (Stages I: local and generally curable, Stage II and III: local and involvement of lymph node, Stage IV Inoperable or metastatic, M0: No distance metastasis, M1:

Table 5.4: Methylation status of *BNC1*, *CCDC8* and *GALNT9* in primary breast tumours from TCGA.

β value of the probes in METASTATIC primary breast tumours from TCGA that are associated with an analysed CpG island promoter region of the genes																
Gene	**Probes**	**M1**	**M2**	**M3**	**M4**	**M5**	**M6**	**M7**	**M8**	**M9**	**M10**	**M11**	**M12**	**M13**	**M14**	**No. of tumours with β value ≥0.70)**
BNC1	cg18560204	0.6441	0.5102	0.2708	0.4231	0.6217	0.5687	0.6609	0.5915	0.3703	0.547	0.099	0.7931	0.6854	0.6424	1
	cg27304406	0.4978	0.3186	0.1607	0.2732	0.4874	0.2298	0.5352	0.6478	0.2044	0.3267	0.0307	0.6101	0.5479	0.4283	0
CCDC8	cg06747432	0.713	0.6201	0.0826	0.1719	0.5959	0.0241	0.6847	0.7195	0.2004	0.0319	0.7821	0.8391	0.6941	0.6542	4
GALNT9	cg02228111	0.1542	0.0347	0.2402	0.219	0.106	0.1097	0.6694	0.0392	0.0526	0.1273	0.0383	0.5663	0.4731	0.088	0
	cg12075445	0.0769	0.0309	0.1638	0.071	0.5271	0.0366	0.7334	0.0296	0.0219	0.1324	0.1901	0.4157	0.356	0.3577	1

Table 5.4: The methylation status of *BNC1* and *GALNT9* in metastatic primary breast tumours from TCGA shows that these three genes are infrequently methylated (β value ≥ 0.70) in metastatic primary tumours, which have metastasised to the distant sites other than the brain such as bone, breast and lungs. CCDC8 is methylated (β value ≥ 0.70) in 4/10 metastatic primary breast tumours from the TCGA. Two of these samples metastasised to liver and bone whereas the metastasis of the other two samples is not known.

BNC1, *CCDC8* and *GALNT9* are infrequently methylated in the 20 non-metastatic tumours from TCGA that were analysed (table 5.5), one of the probes (cg18560204; *BNC1*) is methylated in two tumours and another probe (cg18592647; *BNC1*) is methylated in only of one of the 20 tumours. Similarly, CCDC8, a potential early indicator for metastasising primary breast tumours to the brain, is not methylated in any of the 20 non-metastatic breast tumours analysed from the TCGA. Similarly, *GALNT9* is not methylated in any of the 20 non-metastatic primary tumours analysed.

We carried out functional analyses to investigate the role of *BNC1*, *CCDC8* and *GALNT9* in the metastatic potential of breast cancer cell lines, these are described in chapter 6.

Table: 5.5 Methylation statuses of *BNC1*, *CCDC8* and *GALNT9* in non-metastatic breast tumours from TCGA.

β value of the probes in NON-METASTATIC primary breast tumours from TCGA that are associated with an analysed CpG island promoter region of the genes																					
Gene	**Probe**	**NM1**	**NM2**	**NM3**	**NM4**	**NM5**	**NM6**	**NM7**	**NM8**	**NM9**	**NM10**	**NM11**	**NM12**	**NM13**	**NM14**	**NM15**	**NM16**	**NM17**	**NM18**	**NM19**	**NM20**	**No. of tumours with β value ≥0.70)**
BNC1	cg18560204	0.483	0.609	0.434	0.513	0.148	0.301	0.499	0.485	0.519	0.651	0.61	0.498	0.587	0.442	0.705	0.504	0.714	0.328	0.184	0.561	2
	cg27304406	0.285	0.379	0.224	0.361	0.042	0.261	0.306	0.319	0.277	0.588	0.413	0.108	0.393	0.229	0.56	0.357	0.619	0.335	0.034	0.453	0
CCDC8	cg06747432	0.51	0.67	0.163	0.07	0.581	0.187	0.061	0.383	0.566	0.04	0.456	0.018	0.231	0.521	0.657	0.305	0.148	0.057	0.028	0.374	0
GALNT9	cg02228111	0.027	0.322	0.04	0.043	0.061	0.031	0.39	0.024	0.126	0.155	0.138	0.083	0.323	0.176	0.048	0.476	0.088	0.039	0.046	0.455	0
	cg12075445	0.416	0.018	0.464	0.017	0.142	0.018	0.046	0.016	0.521	0.691	0.27	0.016	0.447	0.024	0.427	0.521	0.025	0.021	0.018	0.589	0

Table 5.5: *BNC1, CCDC8* and *GALNT9* are infrequently methylated in non-metastatic primary breast tumours from TCGA. This data suggests that *BNC1*, *CCDC8* and *GALNT9* are dysregulated specifically in metastatic primary tumours (Methylation: β value ≥ 0.70).

5.3 Discussion

We hypothesized that gene methylation that may contribute to BBM will occur commonly in primary lung tumours, as these often metastasise to the brain in a short time period relative to initial diagnosis (Eichler *et al.*, 2011). Shorter time period of lung tumours to metastasise to the brain may be due to the frequent genomic alterations that give advantages to the lung tumours to invade and/or to proliferate into the brain against the selective pressure of the brain microenvironment. Similarly, breast tumours take a relatively long time to metastasise to the brain compared to the lung tumours (Feld *et al.*, 1984; Zhang *et al.*, 2013). These metastasising breast tumours may have genomic alterations common to lung tumours that are responsible for BBM. In order to identify such novel candidate genes that contributes to BBM, analyses of genome wide 450K methylation array data was carried out.

Our main aim was to identify candidate genes, which are silenced due to promoter hypermethylation and contribute to BBM. Therefore, the 450K-methylation array probes that are not associated with promoter region (those located in gene body and 3' UTR) were discarded. Stringent statistical criteria were used to generate a list of candidate genes from the analyses, which identified six genes *i.e. GALNT9, KRT222, PLEKHA6, TFAP2A, TSPAN4* and *ZNF808*. Methylation analysis of five genes except *KRT222* (which has no well-defined CpG island) in BBM samples identified only *GALNT9* as frequently methylated in 55% (16/29) of BBM samples, and was NOT methylated in any samples from an independent cohort of 33 unrelated primary breast tumours. Moreover, methylation of *GALNT9* correlated to loss of expression of the gene in BBM, suggesting that the methylation of *GALNT9* in BBM contributes to its silencing. *GALNT9* was further investigated for its methylation status in BBM samples

and their corresponding primary tumours from the same patients to investigate if *GALNT9* methylation occurs early or late in the process of tumour evolution.

As described in section 4.3 we predicted that our analysis of unrelated primary breast tumours and BBM would identify two different classes of genes that contribute to BBM. We hypothesized that epigenetic silencing of BBM associated genes would either occur as a) early events (Chambers *et al.*, 2002; Gupta & Massague, 2006) or as b) late events. The dysregulated genes as either involved in: (i) Metastasis initiation (*e.g.* EMT regulators), (ii) Metastasis progression genes or (iii) metastasis virulence gene. Metastatic virulence gene may allow the cancer cells to survive in a foreign tissue environment (Nguyen & Massague, 2007). These are likely to occur as a consequence of the selection pressure provided by the novel environment the metastasised tumour cells find themselves in and as such will be a late event in metastasis evolution *e.g. BNC1* (section 4.3). *GALNT9* may promote BBM by providing selective advantage to metastasised tumour cells in the brain microenvironment. Whereas there may be no advantage in its silencing during the formation of the primary tumour.

GALNT9 encodes a member of the UDP-*N*-acetyl-α-D-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase family of enzymes that catalyze the first step of O-glycosylation; GALNAC-T9 (Chambers *et al.*, 2002). *GALNT9* is expressed most abundantly in the brain and other CNS tissues. It is also expressed, at lower levels, in a number of other tissues including normal breast (Safran *et al.*, 2010).

The GALNAC-T proteins initiate mucin type O-linked glycosylation in the golgi apparatus by the covalent linkage of an α-*N*-acetylgalactosamine (GalNAc) to Ser and

Thr residues (Hanisch, 2001). O-Glycans play an important role in cell adhesion and cell-cell communication and dysregulated glycosylation is a common characteristic of tumour cells (Brockhausen, 2006). A recent screen of GALNAC-T protein expression in a neuroblastoma model, cell lines and tumours found that loss of *GALNT9* expression was linked to a highly malignant phenotype and associated with poor overall and disease free survival (Berois *et al.*, 2013). These findings suggest GALNT9 glycosylates a very specific group of substrates and this may indicate a subtle regulation of transmembrane protein function. Our findings of *GALNT9* promoter methylation, and associated loss of expression, in BBM, but not in primary breast tumours suggest that this proposed subtle change in transmembrane protein function may be a common occurrence in the later stages of the evolution of breast tumour brain metastasis. This is the first time that *GALNT9* has been shown to be dysregulated in cancer by promoter methylation. However, conserved mutations have been identified in approximately 2% of microsatellite instable colorectal cancers (Tuupanen *et al.*, 2014) and *GALNT9* is also mutated, infrequently (<1%), in astrocytoma (Brennan *et al.*, 2013) and lung tumours (Cancer Genome Atlas Research, 2012; Cancer Genome Atlas Research, 2014) and infrequently lost through CNV in breast tumours (Cancer Genome Atlas Research, 2012; TCGA, 2012).

In a small number of samples with clinical information (6 pairs), methylation status of *GALNT9* is not associated with ER/PR/HER2 status of the primary tumours, which have metastasised to the brain in individual patients. The analysis of more samples is needed to determine if *GALNT9* methylation is independent of other clinical features (as in the case in this very small cohort). *GALNT9* was methylated in 4/6 BBM in matched pairs of samples (patient, 1, 2, 8 and 12) whereas it was methylated in only one of the associated primary tumours in individual patients *i.e.* patient 8 (Table 5.6). Patient 2 and

12 are ER+/PR+/HER2- whereas patient 1 is ER+/PR-/HER2-. There was no information on ER+/PR+/HER2- for patient 8. Looking at the tumours stage as well as lymph node and vascular invasion, there is no specific methylation pattern of *GALNT9* that contributes to lymph node or vascular invasion Therefore, it supports the fact that the *GALNT9* is not dysregulated in primary tumours to contribute to the process of invasion/metastasis. However, its dysregulation due its promoter methylation may be crucial for survival of metastasised tumours in the brain (late event).

Table: 5.6 Clinical information of originating primary breast tumours and methylation status *GALNT9*.

Patient	ER status	PR status	HER2 status	Grade and type	lymph/Vascular Invasion	Duration between primary and BBM surgery	Meth in PB	Meth in BBM
Patient 1	Positive	Negative	Negative	Grade I, Invasive Ductal adenocarcinoma,	vascular invasion	5 years		
Patient 2	Positive (5-10%),	Positive (5-10%),	Negative	Invasive lobular carcinoma	NA	10 years		
Patient 5	Positive	Negative	Negative		NA	10 years		
Patient 8	NA	NA	NA	Grade II IDC, Advanced metastatic carcinoma	NA	2 years		
Patient 11	NA	NA	1+ (negative)	Grade III, infiltrating ductal carcinoma	NA	6 years		
Patient 12	Positive	Positive	1+ (negative)	Grade II, infiltrating ductal carcinoma	lymphovascular invasion	5 years		
Patient 15	NA	NA	NA	Grade III, infiltrating Ductal carcinoma	lumphovascular invasion	2 years		

Table 5.6: Clinical information of the primary breast tumours which metastasised to the brain. Methylation status of *GALNT9* in these primary tumours and their corresponding BBM suggests that the methylation status of *GALNT9* may be independent of the clinical characteristics of the primary tumours. None of the primary tumours, which metastasised the brain, are methylated at *GALNT9*.

The methylation status of *BNC1*, *CCDC8* and *GALNT9* in clinical data from TCGA is consistent with our independent analyses of BBM and their associated primary tumours from individual patients. These data support our findings that the *CCDC8* dysregulation occurs early only in those primary breast tumours which may eventually metastasise to the brain where as *BNC1* and *GALNT9* dysregulation occurs late during the process of tumour evolution only after the metastasising tumour cells have left the primary tumour.

5.5 Conclusion:

Analyses of the Infinium BeadChip human methylation 450K array data identified 6 candidate genes that contribute to BBM. We carried out independent experimental validation using CoBRA for 6 candidate genes in BBM samples that identified *GALNT9* as being frequently methylated in BBM tumours. *GALNT9* was not methylated in any of the unrelated primary breast tumours. Furthermore, *GALNT9* methylation correlated to its expression. *GALNT9* was not methylated in any of the primary tumours from the individual patients where as it was methylated in their metastatic counterparts in BBM. This suggests that, the methylation of *GALNT9* occurs late in the process of metastasis. Analysis of the methylation data from the metastatic and non-metastatic tumours from the TCGA was carried out for *BNC1* and *CCDC8* (from candidate gene approach; chapter 4) and *GALNT9*. Analysis of this data supports our findings that the dysregulation of *BNC1* and *GALNT9* due to promoter methylation is a late event whereas the dysregulation of *CCDC8* due to promoter methylation is an early event during the process of BBM. Functional analyses, to determine the contribution of *BNC1*, *CCDC8* and *GALNT9* dysregulation to the metastatic process have been carried out as described in chapter 6.

CHAPTER 6

Functional analyses of *BNC1*, *CCDC8* and *GALNT9* to investigate their influence on the metastatic potential breast cancer cell lines.

6.1 Introduction

Invasion and metastasis, two important hallmarks of cancer (Hanahan & Weinberg, 2000), are associated with various complex genetic and epigenetic aberrations allowing tumours to disseminate to distant sites through the lymphatic and vascular systems (Barekati *et al.*, 2012). The invasion of a tumour from lymphatic and blood vessels giveS rise to an eventual colonization of the tumour at a suitable niche and to grow into micro (Leong *et al.*, 2011) and macro metastases (Barekati *et al.*, 2012; Kurbasic *et al.*, 2015). It is important to understand and to identify those genetic and epigenetic alterations in primary tumours (such as breast) in order to provide improved prognostic, diagnostic and management of metastatic tumours. In the case of primary breast tumours, lymph nodes metastases are taken as an important prognostic indicator to evaluate the possibility of distance metastases. As recent study has reported epigenetic aberrations status of 12 genes in primary breast tumours and the lymph node metastases including aberrations in DNA methylation status of *BMP6*, *BRCA1* and *P16*, which could be used as prognostic markers in breast cancer metastases to lymph nodes (Barekati *et al.*, 2012). However, epigenetic mechanisms underlying distance metastases of breast tumours remain very poor.

Previous studies have reported dysregulated genes in breast cancer metastases to lungs

and bone (Kang *et al.*, 2003; Minn *et al.*, 2005a) and the genes with reduced mRNA expression in breast cancer brain metastases (Stark *et al.*, 2005) (see 1.6.4 for details). This implies that the metastatic process may involve a group of genes dysregulated in selectable subpopulations of primary breast tumours that offer these cells an advantage to metastasise to the specific distant site and evolve into metastatic tumours. Consistent with this, three potential candidate metastatic suppressor genes; *BNC1*, *CCDC8*, and *GALNT9*, which are dysregulated either both in primary and BBM (*CCDC8*) and or in brain metastases (*BNC1* and *GALNT9*) may provide an advantage to the metastasising cells to invade the brain or to grow and proliferate within the intracranial environment.

Cancer cells need to migrate through the blood or lymph to reach a distant site during the process of metastasis. Therefore, we have employed *in vitro* cell migration and invasion assays (Kramer *et al.*, 2013) to determine if the silencing of the identified genes may contribute to the metastatic process in breast cancers.. We have employed a wound-healing assay (a scratch assay) to measure the capacity of cells in a loosely connected population to move in any direction without obstructive fiber networks towards the wound created. Invasion, a movement of cells through ECM barrier either by modifying its own shape or by proteolytic degradation of ECM network, is a crucial step in the process of metastasis (Kramer *et al.*, 2013). Therefore, invasion requires ECM remodeling, as cells need to undergo adhesion, proteolytic degradation of ECM components and then migration during the process of intravasation and extravasation to metastasise into a distant niche. In this regard, migration, a non-destructive and non-proteolytic movement is a prerequisite for invasion as the cells are able to migrate without invasion (Kramer *et al.*, 2013). Therefore, we carried out an invasion assay only if the cells exhibited the potential to migrate following the knockdown of *BNC1*, *CCDC8* or *GALNT9*.

6.2 Result

6.2.1 Treatment of breast cancer cell lines with 5-AZA-2-dC

We treated the breast cancer cell lines MCF7, T47D, MDA-MB231, BT549 and ZR75 with the demethylating agent; 5-AZA-2-dC to investigate the reexpression of *BNC1*, *CCDC8*, and *GALNT9* in these cell lines. Prior to treatment, methylation analyses of the genes were carried out on these cell lines. *BNC1* was methylated (appendix D1) and downregulated in MCF7 (figure 6.1), where as it was methylated and silenced in T47D, MDA-MB231, BT549 and ZR75. *BNC1* showed an increased level of expression in MCF7 and it was reexpressed in T47D and ZR75, after treatment with 5-AZA-2-dC. However, there was no evidence of *BNC1* expression in MDA-MD231 and BT549. Similarly, *CCDC8* (appendix D1) was methylated in all the cell lines treated. *CCDC8* was downregulated in MCF7 (figure 6.1) whereas it was silenced in other cell lines (T47D, MDA-MD231, BT549 and ZR75). Following treatment with 5-AZA-2-dC, *CCDC8* was up regulated in MCF7 and was reexpressed fully in ZR75. *CCDC8* did not re-express in MDA-MB231 and BT549 and was expressed in T47D prior to treatment, and the expression level has persisted after the treatment with 5-AZA-2-dC. *GALNT9* (figure 6.1) was downregulated in MCF7, and was silenced in T47D, whereas it was expressed in MDA-MD231 and BT549. *GALNT9* was up regulated in ZR75 after the treatment. However, there was no reexpression or upregulation in T47D or MCF7.

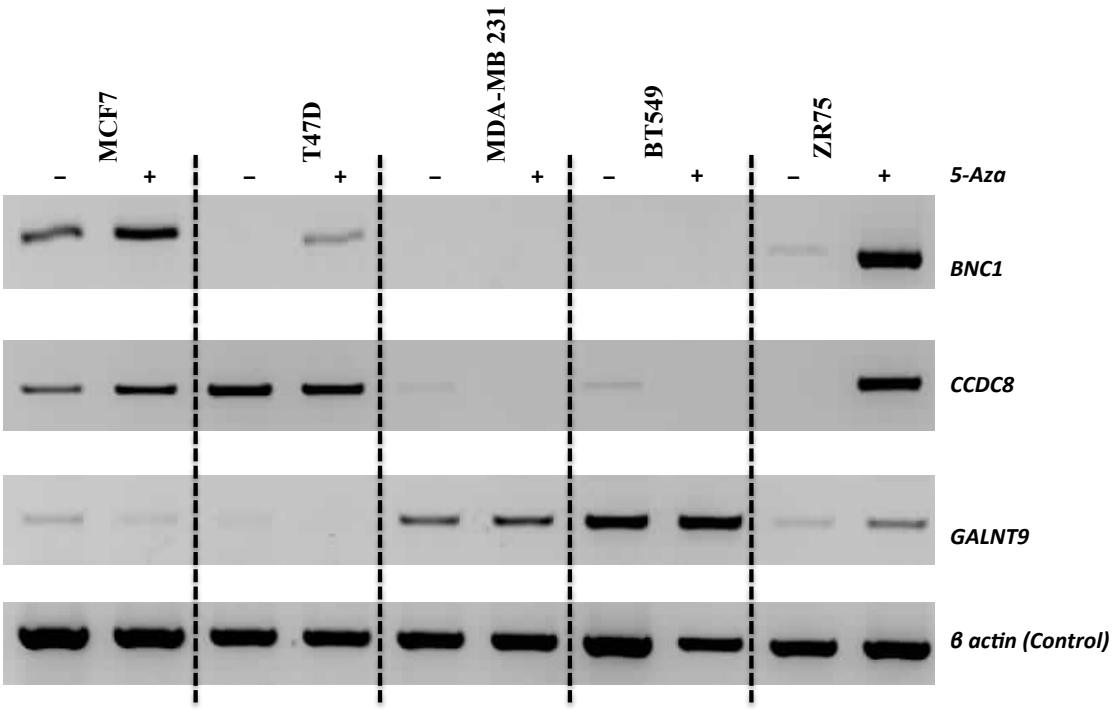


Figure 6.1: Global demethylation resulted in the reexpression of *BNC1, CCDC8* and *GALNT9* in breast cancer cell lines. Reverse transcription PCR (RT-PCR) showed that treatment of breast cancer cell lines with 5-AZA-2-dC, a demethylating agent resulted in re-expression of *GALNT9, CCDC8* or *BNC1* in breast cancer cell lines. For comparison, endogenous expression is also shown (-) along with the expression after the global demethylation by 5'aza treatment (+). β actin was used as a control. The cell lines where *BNC1*, *CCDC8* and *GALNT9* are expressed *i.e. BNC1* in MCF7, *CCDC8* in T47D and, *GALNT9* in MDA-MB231 and BT549 were used in our *in vitro* knockdown experiments. The experiment was carried out in three replicates.

6.2.2 Loss of *GALNT9*, *CCDC8* or *BNC1* expression increases the motility of breast cancer cell lines.

BNC1, *CCDC8* and *GALNT9* were knocked down by RNA interference (using siRNA oligos) to study the migratory behavior of breast cancer cell lines. *BNC1* was knocked down in MCF7, *CCDC8* was knocked down in T47D and, *GALNT9* was knocked down in MDA-MB231 and BT549. The knockdown for these genes was carried out as described in section 3.6.2. Before carrying out the migration assay for the genes, the knockdown efficiency of each siRNA oligo was determined by RT-PCR (Figure 6.2). After determining knockdown efficiency of *BNC1*, *CCDC8* and *GALNT9* individually, migration assay was set up in triplicates to minimize experimental biases. An identical transfection experiment was set up in parallel to ensure the knock-down efficiency at the time of migration assay experiments, which were then confirmed by RT-PCR and Western blot (Figure 6.5).

Forty-eight hours after initial transfection with siRNA oligos against *BNC1*, *CCDC8* or *GALNT9* breast cancer cell lines showed loss of specific gene expression. Control oligos was used as a control in a respective cell line (Figure 6.2 and 6.3)

Following knockdown, cells were seeded in serum-free media to form confluent monolayers. The monolayers were scratched with a 200µl pipette tip to generate "Wounds". Photographs were captured at the time of scratch (labeled as a zero hour) and at 48 hours after the initial scratch at 100X magnification using a digital camera attached with the microscope. The widths of each scratch were measured at three separate points to generate an average distance using Image J software (Schneider *et al.*, 2012). The difference in the migratory distance at zero and 48 hours was calculated and

was compared with a cell line transfected with scrambled control oligos. It was observed that the knockdown of any of these three genes increased the migratory potential compared to cell lines transfected with control oligos. The increase in motility of cell lines following knockdown of *BNC1* (Figure 6.3; B), *CCDC8* (Figure 6.3; A) or *GALNT9* (Figure 6.3; C) was statistically significant (*BNC1*: p=0.011, *CCDC8*: p=0.001, *GALNT9*: p= 0.027) compared to control cells (scrambled siRNA transfected).

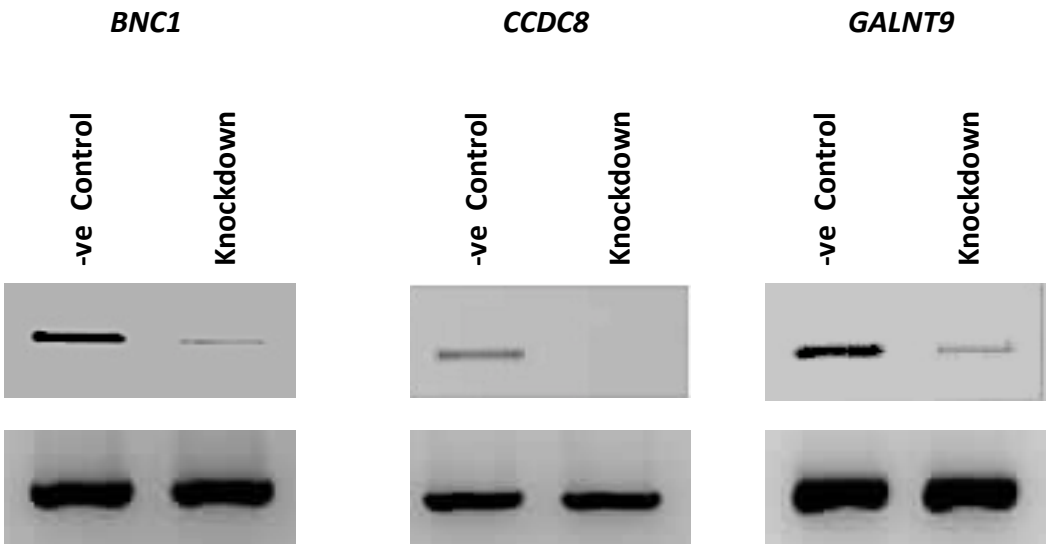


Figure 6.2: Reverse Transcription (RT) PCR to confirm Knockdown of *BNC1*, *CCDC8*, and *GALNT9* in breast cancer cell lines prior to functional assays. RT-PCR products of *BNC1, CCDC8,* and *GALNT9* transcripts in breast cancer cell lines (MDA-MD231, T47D and MCF7 respectively after siRNA knockdown compared to transfection with a control siRNA (upper panel). Equal loading was confirmed by analysis of *β actin* expression (lower panel).

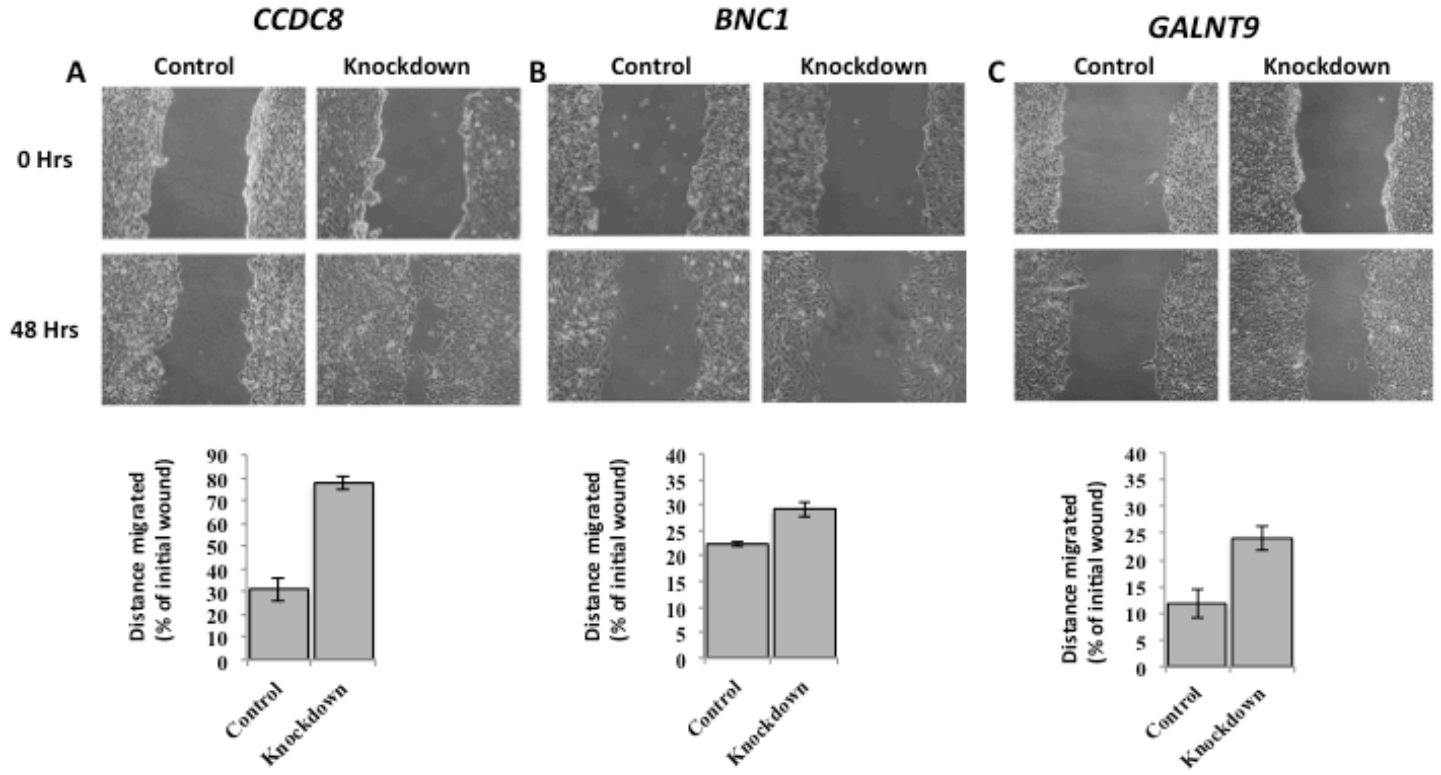


Figure 6.3: Loss of *CCDC8, BNC1* or *GALNT9* expression increases the migratory potential of breast cancer cell lines. The breast cancer cell lines that expressed *CCDC8, BNC1* or *GALNT9* were identified; these genes were knocked down by siRNA (see figures 6.2 and 6.5). (A) T47D cell lines transfected with siRNA oligos against *CCDC8;* (B) MCF7 cell lines transfected with siRNA oligos against *BNC1* or (C) MDA-MB231 cell lines transfected with siRNA against *GALNT9* exhibited more migratory potential compared to respective cell lines transfected with control siRNA oligos. Following siRNA transfection, confluent cells were incubated in serum-free media and an artificial wound was scratched through them (0 hrs). 48hrs later the distance migrated was

calculated by subtracting the value of non-migrated distance from the initial wound. The distance migrated (in percentage), by respective cell lines knocked down with siRNA against *CCDC8, BNC1,* or *GALNT9* compared to the respective cell lines transfected with control siRNA oligos, was statistically significant (p=0.001, 0.011 and 0.027 respectively).

6.2.3 Reduced expression of *BNC1*, *CCDC8* and *GALNT9* increases the invasive potential of breast cancer cell lines

BNC1, *CCDC8* and *GALNT9* were knocked down in breast cancer cell lines by siRNA oligos as described in section 3.6.2. The respective cell lines that were knocked down with *BNC1*, *CCDC8* or *GALNT9* were applied to matrigel-coated invasion chambers with 8-μm polyethylene terephtphalate membrane pores. Forty-eight hours later, cells that had "invaded" through the pores were stained with crystal violet and photographs were taken at 100X magnification.

The cells were isolated and the OD was measured at 540nm in order to quantify the proportion of cells that had invaded through the matrigel. The number of breast cancer cell that invaded through the matrigel coated pores following *BNC1* knocked down was increased by 40% (p=0.006) compared to cells transfected with control oligos. (Figure 6.4A). Similarly, following knockdown of *CCDC8,* 27% more cells invaded through the matrigel compared to cells transfected with control oligos (p=0.021) (Figure 6.4B). Following knockdown of *GALNT9* 35% more cells invaded (p=0.025) compared to cell transfected with the control-scrambled siRNA (Figure 6.4C).

Increased motility and invasive potential following reduction of expression of these genes suggests that these candidates may be involved in the regulation of normal cellular physiology and that loss of their expression may contribute the metastatic process in breast tumour cells.

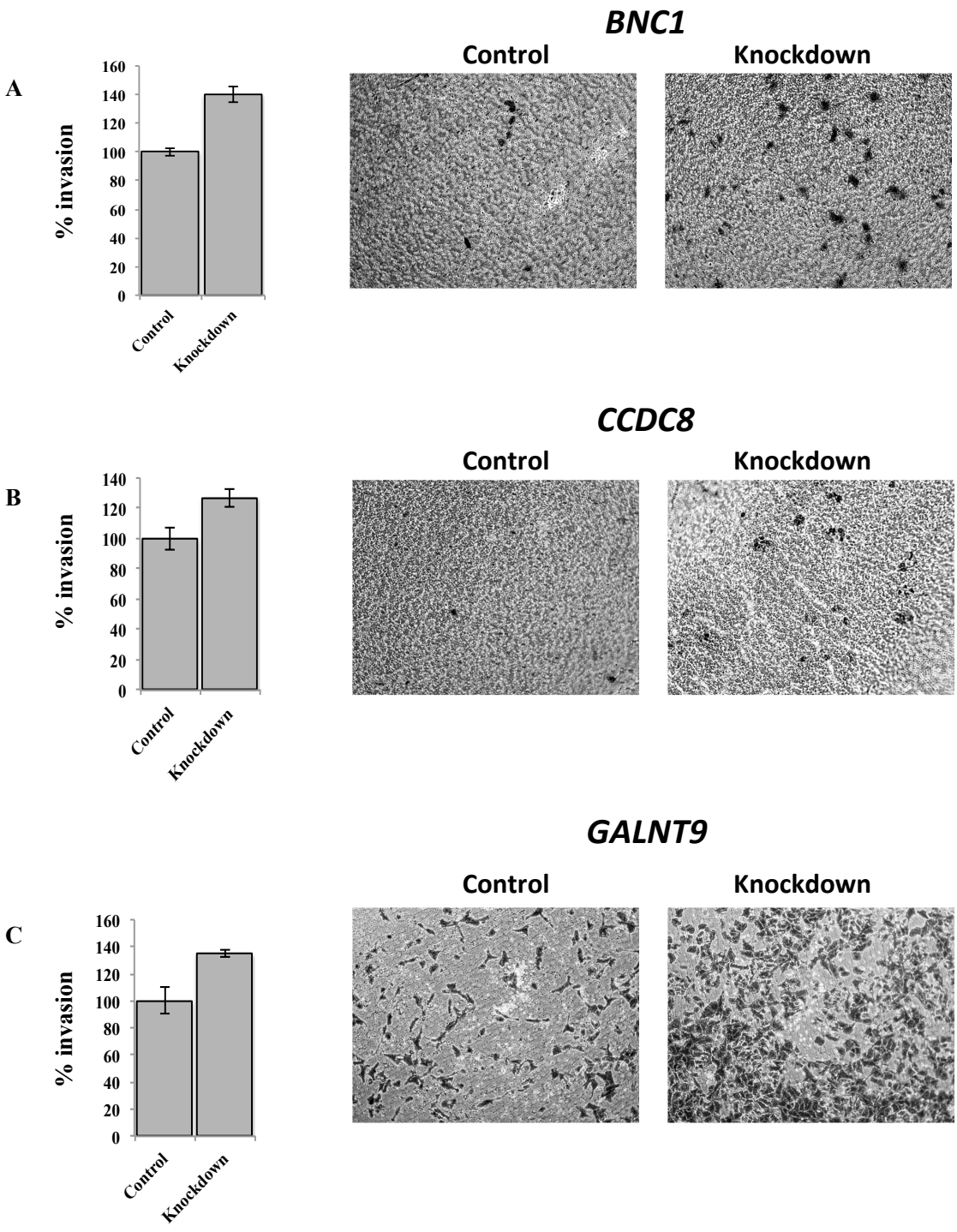


Figure 6.4: Reduced expression of *BNC1, CCDC8* or *GALNT9* increases the invasive potential of breast cancer cell lines. A trans-well invasion assay was carried out following the knockdown of *BNC1*, *CCDC8* or *GALNT9* in breast cancer cell lines. The invasive capacity of these cells was compared with the same cell lines transfected with scrambled control siRNA oligos (control). The numbers of cells that had invaded a matrigel coated micropore membrane was determined

colourimetrically 48hrs after initial seeding. (A) MCF7 cell lines transfected with siRNA oligos against *BNC1*, (B) T47D cell lines transfected with siRNA oligos against *CCDC8* and (C) MDA-MB231 cell lines transfected with siRNA oligos against *GALNT9*, exhibited a statistically significant increase in invasiveness compared to negative control siRNA transfected cells. p=0.001 (*BNC1*), p=0.021 (*CCDC8*) and p=0.025 (*GALNT9*), Invasive potential was calculated as a percentage increase above that observed for the control cells (% invasion).

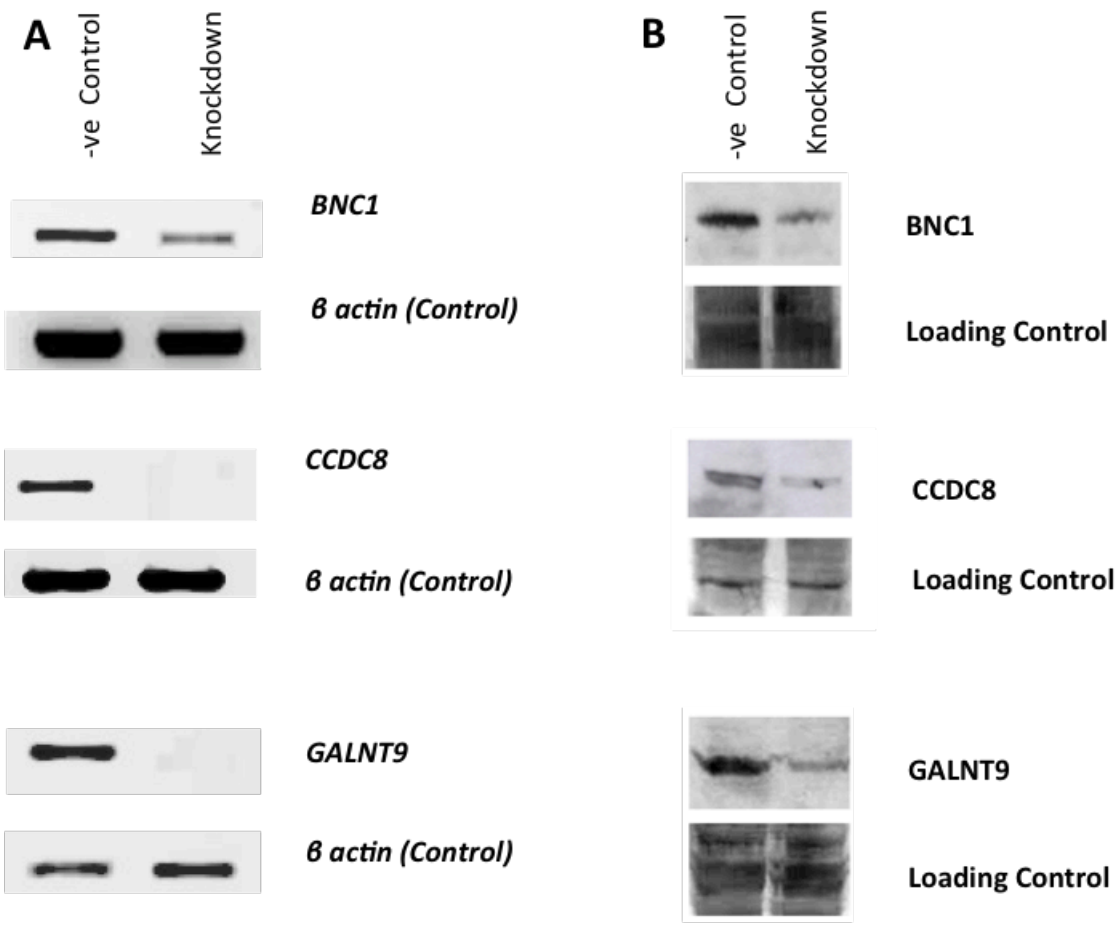


Figure 6.5: Knockdown of *BNC1*, *CCDC8,* and *GALNT9* in breast cancer cell lines set up in parallel during migration and invasion assays is confirmed by Reverse Transcription (RT) PCR and Western blot. (A) RT-PCR products of *BNC1, CCDC8,* and *GALNT9* transcripts in breast cancer cell lines (MDA-MD231, T47D and MCF7 respectively), after siRNA knockdown compared to transfection with a control siRNA and (B) western blot of GALNT9, CCDC8, and BNC1 proteins to confirm their knockdown in each respective cell line. 70 μg of protein was loaded in each lane. Equal loading was confirmed by β actin for RT-PCR and by staining total protein with India ink for western blot.

6.2.4 Association of reduced expression of BNC1, *GALNT9* or *CCDC8* with poor relapse-free survival of patients.

The clinical significance of the expression of *BNC1*, *CCDC8* and *GALNT9* was analysed using publically available GEO expression profiles using the prognoscan database (Mizuno *et al.*, 2009). Given a patient population with varying disease outcomes, Prognoscan partitions that population into a high-expression and low-expression group for each gene by choosing a threshold that maximizes the statistical significance of difference in outcome. It also corrects for multiple testing using the method of Miller and Siegmund (Miller & Siegmund, 1982). This analysis demonstrated that, in two independent datasets, low *CCDC8* expression was significantly associated with poor relapse free survival (GSE12276: p=0.001, GSE1456-GPL97: p=0.004) (Figure 6.6B) and in one data set, low *GALNT9* expression was associated with poor relapse free survival, (GSE1379: p=0.0029) (Figure 6.6A). There was no evidence in any of the datasets analysed that low *BNC1* expression correlated with poor relapse free survival or any other clinical indicator (Figure 6.6C). The survival curves associated with *BNC1*, *CCDC8* and *GALNT9* in different datasets are given in appendix D2.

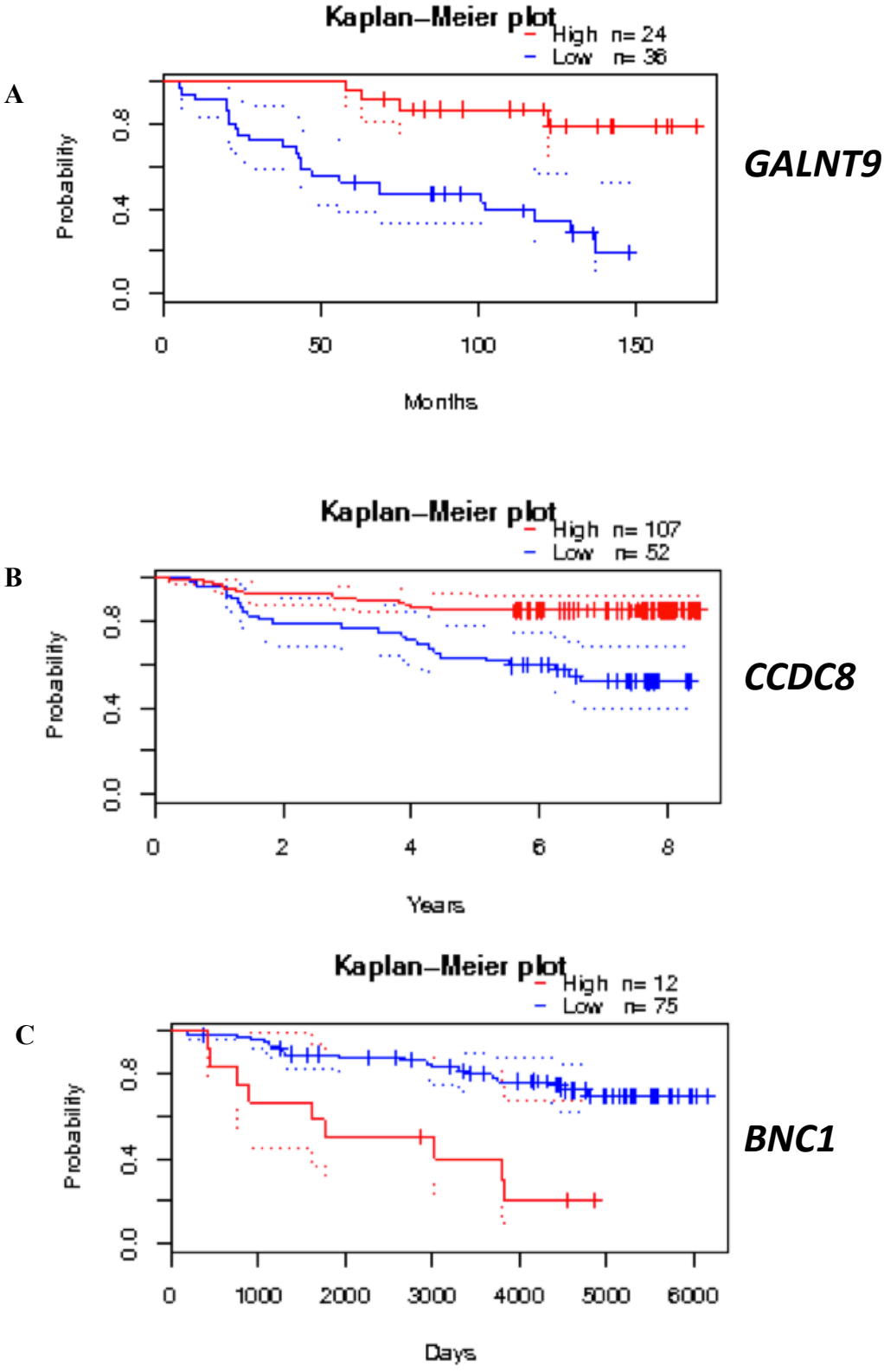


Figure 6.6: Loss of expression of *CCDC8* and *GALNT9* correlates with relapse-free patient survival. Kaplan-Meier analysis of multiple gene expression studies via the prognoscan database revealed that in two separate studies, low expression of (A)

GALNT9 **(p=0.003) and (B)** ***CCDC8*** **(p=0.001) is associated with poor relapse free survival. In contrast to our findings, (C) expression of** ***BNC1*** **is associated with relapse-free survival of patients.**

6.3 Discussion

Metastasis to the brain is an increasingly common event in the progression of breast cancer. Given the extremely poor clinical outcome following a diagnosis of BBM (Sperduto *et al.*, 2010), it is imperative that the underlying molecular biology that drives tumour evolution to the colonization of the brain is revealed. An improved understanding of these events will identify novel therapeutic targets and prognostic markers.

To date some progress has been made to identify prognostic markers for breast cancer metastasis by gene expression profiling (Paik *et al.*, 2004) (See section 1.6.5). However, prediction of site specific-metastasis and a late recurrence, a common feature of BBM, or indeed any site-specific recurrence has proven difficult to predict (Weigelt *et al.*, 2005; Burstein & Griggs, 2012) (see section 1.6.5).

As described in section 4.2, it is possible that those genes infrequently methylated in primary breast tumours and frequently methylated in BBM samples may be involved in BBM. In this regard, loss of *CCDC8* expression (*CCDC8* is commonly methylated both in primary tumours and BBMs in the same patient) may provide an advantage to primary tumours; contributing to invasion into and to proliferation within the brain (an early event). Similarly, *BNC1* and *GALNT9* are unmethylated in primary tumours and frequently methylated in BBM samples from the same patient. Loss of expression of these genes may provide survival advantages to tumour cells in the brain (late events).

Our functional analyses included *in vitro* models to assess migratory and invasive potential of these genes in breast cancer cell lines. Our wound healing assays on respective breast cancer cell lines knocked-down with *BNC1*, *CCDC8* or *GALNT9*

showed that these genes increase the migratory potential of the cells significantly compared to control cell lines (figure 6.3). Similarly, invasion assays of respective breast cancer cell lines knocked down with *BNC1*, *CCDC8* or *GALNT9* showed significant increases in invasive potential (Figure 6.4).

Kaplan-Meier (K-M) analysis using the Prognoscan databases has been used to investigate the survival curves i.e. overall survival (OS) and disease free survival (DFS) of the patients due to the downregulation of *BNC1*, *CCDC8* or *GALNT9*. OS *i.e.* the time a patient survives after certain date of event such as surgery or a particular treatment with/out any disease or disease recurrence (Chua, 2005) or RFS *i.e* the time a patient survives without having any disease or a disease recurrence (after the disease has been fully removed by a surgery and treated with adjuvant therapy) (Chua, 2005) are caulcated using K-M analysis. DFS is generally calculated in every three years, which then serves as a secondary point for OS, which is calculated in five years durations if patients alive. Kaplan-Meier analyses demonstrated that low *CCDC8* and *GALNT9* expression was significantly associated with poor relapse free survival of the patients. This implies that downregulation or silencing of these two genes contributes to disease recurrence or the expression of these two genes is important to suppress the chances of tumours to recur. In contrast to our findings, there was no evidence that low *BNC1* expression correlated with poor relapse free survival or any other clinical indicator.

Basonuclin 1 (BNC1) is a zinc finger transcription factor that interacts with the promoters of both RNA polymerases II and I (Zhang *et al.*, 2007). BNC1 target genes have been implicated in a broad range of functions (Ma *et al.*, 2006; Wang *et al.*, 2006; Zhang *et al.*, 2007). *BNC1* knockdown in mouse oocytes shows that BNC1 may be

necessary for embryogenesis and oogenesis (Ma *et al.*, 2006; Vanhoutteghem *et al.*, 2009) (see section 4.3 for details of BNC1 functions).

Analysis of HumanMethylation 27K and 450K array data from TCGA indicates that *BNC1* Promoter methylation is an infrequent event in both metastatic and non-metastatic primary breast tumours (TCGA, 2012) (Section 5.3.1). Expression of BNC1 is induced by transforming growth factor-β1 (TGFβ1) signalling and, in turn, it acts as a transcription factor for a number of modulators of epithelial dedifferentiation during the process of EMT (Feuerborn *et al.*, 2014) (Section 4.3)**.** BNC1 may be a component of a network of transcription factors that influences epithelial cell plasticity as well as TGF-β1 signaling and its silencing may contribute to EMT influencing invasion and metastasis (Feuerborn *et al*., 2014). These findings suggest that the expression of BNC1 would enhance the process of metastasis via EMT. Our findings are consistent with this; we find that *BNC1* is infrequently methylated in primary breast tumours (17%) and frequently methylated and silenced in BBMs (73%). In addition, we have shown that *BNC1* promoter methylation is a late event in tumour evolution, only occurring in the brain metastasis of a BBM patient and not in the associated primary tumour. It is plausible that *BNC1* expression is commonly required for EMT to occur during metastasis and, once these cells have metastasised to the brain, loss of *BNC1* expression contributes to Mesenchymal to epithelial transition (MET).

It is important to note that the *BNC1* overexpression did not exhibit any differences in migration and invasion when it was transfected in pancreatic cell line (Yi *et al.*, 2013). According to our findings, *BNC1* silencing is a late event during the process BBM. This implies that *BNC1* contributes to survival of metastasized tumours to the brain. Our *in vitro* wound healing and invasion assays have shown that *BNC1* knockdown with RNAi

significantly increased the migratory and invasive potential of breast cancer cell lines. This suggests that the *BNC1* downregulation or silencing may contribute to metastasis of breast tumours to other distant sites than the brain, however, in BBM, dysregulation of *BNC1* occurs late, its dysregulation may be an important factor for tumour cells to survive against the selective pressure of the brain microenvironment. In addition, *BNC1* silencing may contribute to local invasion of the tumours inside the brain.

CCDC8 encodes a coiled-coil domain containing protein (CCDC8) that is one of three proteins that are, mutually exclusively, mutated in patients with 3M-syndrome (Hanson *et al.*, 2011). (Hanson *et al.*, 2011) (Section 4.3). It has been shown that loss of any 3M complex protein significantly altered the interphase microtubule network (Yan *et al.*, 2014). The core 3M-protein complex interacts with CUL9, which has been proposed to mediate the functions of the 3M-complex via the ubiquitylation and degradation of survivin (Li *et al.*, 2014). The 3M-complex also interacts with the F box protein FBXW8, ROC1 and the tumour suppressor p53 (Yan *et al.*, 2014) suggesting it may contribute to correct cellular physiology through multiple mechanisms. Taken together, these discoveries suggest that CCDC8 downregulation or silencing contributes to microtubule damage suggesting its role on microtubule integrity (Yan *et al.*, 2014). Microtubules are an important cytoskeleton component that is crucial for spindle fibres formation and cytokinesis during mitosis (Yan *et al.*, 2014). Our findings have shown that the downregulation or silencing of CCDC8 increases the metastatic (migratory and invasive) potential of the cells, which could possible that the silencing of *CCDC8* may disintegrate microtubule giving an advantage to cancer cells to migrate and invade through ECM contributing to metastases.

CCDC8 interacts with Tip60, a protein required for acetylation of p53 and promotes p53 acetylation thereby increasing p53 mediated apoptotic response (Dai *et al.*, 2011). Therefore, CCDC8 is required specifically in p53 mediated apoptotic process following DNA damage and its downregulation reduces the apoptotic response mediated by p53 (Murray *et al.*, 2013). It is possible that a decreased apoptotic response mediated by P53 due to CCDC8 silencing may influence cellular physiology giving selective advantage to cancer cells contributing to their capacity to proliferate and to invade the tissue promoting metastases.

GALNT9 encodes a member of the UDP-*N*-acetyl-α-D-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase family of enzymes that catalyze the first step of O-glycosylation; GALNAC-T9 (Toba *et al.*, 2000). The first step of mucin type O linked glycosylation is the transfer of a α-*N*-acetylgalactosamine (GalNAc) residue to the hydroxyl group of Ser or Thr in target protein in the golgi apparatus (Hanisch, 2001; Peng, 2010). GALNT9 is expressed most abundantly in the brain and other CNS tissues. It is also expressed, at lower levels, in a number of other tissues including normal breast (Safran *et al.*, 2010).

O-Glycans play an important role in cell adhesion and cell-cell communication and dysregulated glycosylation is a common characteristic of tumour cells (Brockhausen, 2006). Mucin 1 (MUC1), in particular, has been identified as a highly O-glycosylated transmembrane protein that is dysregulated at the expression and post-translational level in multiple tumour types (Brockhausen, 2006). MUC1 is commonly overexpressed but under-glycosylated in primary breast tumours (Hayes *et al.*, 1991; Perey *et al.*, 1992) and the expression of under-glycosylated MUC1 is associated with high tumour grade,

metastatic potential and invasiveness of breast tumours (de Roos *et al.*, 2007; van der Vegt *et al.*, 2007; Ghosh *et al.*, 2013) (Section 5.3).

GALNT9 is a member of a sub family (with GALNT8, 18 and 19) that differ significantly in sequence from other GALNAC-T members (Raman *et al.*, 2012) and as such does not have catalytic activity towards classic MUC1 variants derivatives (MUC1/A, MUC5AC or mono-GalNAc/Thr7thEA2) (Li *et al.*, 2012). Moreover, GALNT9 is capable of transferring GalNAc to a very specific subset of synthetic peptides (see details on section 5.3) (Zhang *et al.*, 2003). These findings suggest GALNT9 glycosylates a very specific group of substrates and this may indicate a subtle regulation of transmembrane protein function. Our findings of *GALNT9* promoter methylation, and associated loss of expression in BBM, but not in primary breast tumours suggest that this proposed subtle change in transmembrane protein function may be a common occurrence in the later stages of the evolution of breast tumour brain metastasis. Hence, changes in simple mucin type-O-glycans could give an advantage to metastasised tumour cells to differentiate and to proliferate into a macrometastases.

This is the first time that *GALNT9* has been shown to be dysregulated in cancer by promoter methylation. However, conserved mutations have been identified in approximately 2% of microsatellite instable colorectal cancers (Tuupanen *et al.*, 2014) and *GALNT9* is also mutated, infrequently (<1%), in astrocytoma (Brennan *et al.*, 2013) and lung tumours (Cancer Genome Atlas Research, 2012; Cancer Genome Atlas Research, 2014) and infrequently lost through CNV in breast tumours (Cancer Genome Atlas Research, 2012; TCGA, 2012).

6.4 Conclusion

Our functional analyses to investigate the role of *BNC1*, *CCDC8* and *GALNT9* in BBM presented here, indicate that silencing of these in breast tumours may contribute to metastasis to the brain and possibly other distant organs. There was a significant increase in cell motility and invasiveness following knockdown of *BNC1*, *CCDC8*, *and GALNT9* in breast cancer cell lines. Downregulation of *CCDC8* and *GALNT9* is associated with poor relapse-free survival of breast cancer patients. Further studies will be required to determine if epigenetic changes/dysregulation of these genes are detectable in circulating tumour cells, micrometastases, in other reservoir sites or only in macroscopic brain metastases. Further functional assessment should be carried out on these genes to understand the cellular function and their specific roles on metastases. However, our current level of knowledge about these genes suggests that their dysregulation is functionally significant and may provide useful prognostic markers for BBM and be novel therapeutic targets.

CHAPTER 7

Genome-wide 450K methylation analyses to identify genes dysregulated in BBM

7.1 Introduction

One of the keys to identifying prognostic markers specific to BBM is to identify genes that are uniquely dysregulated in breast tumours metastasising to the brain. These dysregulated genes may either be metastatic promoter genes or be metastatic suppressor genes contributing to BBM. The metastatic suppressor genes are infrequently methylated or unmethylated in primary breast tumours and frequently methylated in BBM whereas the metastatic promoter genes are expected to be frequently methylated in primary breast tumours and are subsequently demethylated or hypomethylated in BBM. In both cases, the differential methylation of genes may be attained by comparison of the methylation status of the genes in primary breast tumours and BBM samples.

Our previous strategies in screening genes included identifying genes which were methylated in one of the tumour types that readily metastasised to the brain (section 4) or to identify those probes (individual CpG sites), that are differentially methylated between primary breast and lung tumours by statistical analyses of the 450K methylation data from the TCGA (section 5). Those genes that met the criteria were experimentally validated in BBM samples and unrelated primary tumours from the patients.

Only 2% of the human genome consists of protein coding regions. Very little information is currently available on the role of non-coding and regulatory regions in cancer (Weinhold *et al.*, 2014). It is crucial to explore the non-coding component of the cancer genome to identify regulatory regions and other novel targets to improve prognosis and clinical outcomes (Weinhold *et al.*, 2014).

To overcome the limitations of candidate gene approaches, we carried out genome-wide screening of methylation status in BBM. We used 24 BBM samples from the patients to carry out Infinium BeadChip 450K methylation array, which is one of the most accurate, comprehensive and highly reproducible technologies available today (Dedeurwaerder *et al.*, 2011; Fortin *et al.*, 2014). 450K methylation array takes an account of the CpGs located both in CpG-rich islands as well as in non CpG-rich sites (Zhang *et al.*, 2012). This technology makes it possible to map individual probes (CpG dinucleotides) with differential methylation levels (Bibikova *et al.*, 2009) in primary tumours and BBM. This analysis allows both the identification of CpG hypermethylation of candidate tumour/metastasis suppressor genes and CpG hypomethylation of candidate proto-oncogenes/metastasis promoting genes in BBM (Dedeurwaerder *et al.*, 2011).

7.2 Results:

7.2.1. Normalisation of the TCGA data sets

Before comparing the normal breast tissues, primary breast tumours and BBM samples at a genome wide level, normalization of the TCGA methylation data was carried (in the same manner as our metastatic tumour data was normalized) out to avoid unwanted differences (Fortin *et al.*, 2014) in the array data sets. Normalisation was carried out Dr.

David Huen, a bioinformatics lecturer at School of Biology, Chemistry and Forensic Sciences, Faculty of Science of Engineering, University of Wolverhampton.

7.2.2. Statistical analyses of the 450K-methylation data to identify genes differentially methylated in normal breast tissues and primary breast tumours, and BBM samples

Genome wide 450K methylation data from the TCGA was downloaded for 14 normal breast tissues (all that were available) and 20 primary breast tumours (see appendix E1 for TCGA barcoded identifiers). After the normalization for these data was carried out (see section 7.2.1), two independent approaches were used to identify the probes, which were differentially methylated between the primary breast tumours and BBM samples. As a first approach, an average of the beta value of each probe in all the samples was calculated in primary breast tumours and in BBM. Those probes, which had a mean differential methylation between primary tumours and BBM, were identified. This approach identified 56 genes, of these only 29 genes had differentially methylated probes in their promoter region (Appendix E2). Genes where methylated probes were not in promoter regions were discarded from further analysis. Of 29 genes, only 9 genes consisted significant level of differences in their beta values between primary tumours and BBM. These genes are: *HSPB9*, *MIR1179*, *LOC154872*, *EDARADD*, *DDX52*, *RNF8, SOX5*, *RBM23* and *LMX1B* (table 7.1). However, two of these genes *i.e. MIR1179* and *LOC15472* did not have well-defined upstream CpG islands. Therefore, it was decided not to carry these two genes forward in the analysis. The methylation status of the remaining 7 genes (*HSPB9*, *EDARADD*, *DDX52*, *RNF8*, *SOX5*, *RBM23* and *LMX1B*) was determined in 15 BBM samples. Only one gene, *HSPB9* was frequently methylated (86%) in BBM samples, and none of the other six genes were methylated in

any of the BBM samples (table 7.2, Figure 7.1A). Therefore, only *HSPB9* was carried forward to determine its methylation status in an unrelated cohort of 20 primary breast tumours. However, *HSPB9* was also frequently methylated (89%) in primary tumours (Table 7.2, figure 7.1B). Taken together, none of the genes from the first approach were significantly differentially methylated.

Table 7.1: Genes identified from the 450K-methylation data analyses based on mean β value of all the samples.

Gene Symbol	Entrez ID	Chromosome	Mean (β value) BBM	Mean (β value) primary tumours	Difference
HSPB9	94086	chr17	2.029788864	0.0868493	2.130872037
DDX52	11056	chr17	1.525909178	0.417176197	1.118207247
EDARADD	128178	chr1	0.422072021	0.726856306	1.192213701
RNF8	157360	chr6	0.822653313	0.524739581	0.297913731
SOX5	281485	chr12	0.837062213	0.556376304	0.28068591
RBM23	311885	chr14	0.818144254	0.538907402	0.279236851
LMX1B	224781	chr9	0.826886808	0.54627828	0.280608528

Table 7.1: 450K methylation data of the BBM samples and primary tumours from the TCGA was analysed to identify probes having differential methylation status in BBM compared to primary breast tumours. The mean β value of all the samples for each probe was used as a basis to find the differences in the methylation in primary and BBM samples. This analysis resulted in 56 probes (genes), of which 29 probes were located in the promoter regions of the respective gene (appendix E2). Only 7 of these probes (genes) were selected based on the differences in the β value (greater than 0.25), and the presence of well-defined CpG Island for experimental validation of methylation status of these genes using CoBRA (Note, β-values have been renormalized for comparison and thus do not sit within the range of 0-1).

Table 7.2: Methylation status of the genes in BBM and in a cohort of unrelated primary breast tumour samples

Gene	Breast to Brain metastases (BBM) samples (n=15)															% Meth (n=15)
	BM1	BM2	BM3	BM4	BM5	BM6	BM7	BM8	BM9	BM10	BM11	BM12	BM13	BM14	BM15	
HSPB9																86
DDX52																0
EDARADD																0
SOX5																0
RBM23																0
RNF8																0
LMX1B																0

Gene	Methylation status of HSPB9 in a cohort of unrelated primary breast tumour samples (n=19)																			% Meth (n=19))
	BP136	BP137	BP138	BP139	BP140	BP141	BP142	BP143	BP144	BP145	BP146	BP147	BP148	BP149	BP150	BP151	BP153	BP167	BP170	
HSPB9																				89

Table 7.2: Promoter methylation status of seven candidate genes in a cohort of BBM and unrelated primary breast tumour samples. These genes were identified from the 450K-methylation array data analysis based on the differences between mean methylation (β value) in BBM compared to primary breast tumours from the TCGA. 15 BBM samples were used to investigate the methylation

status of these genes using CoBRA. Only *HSPB9* was frequently methylated in these BBM samples (86%), which was further analysed in primary tumours and found to be methylated in 89% of the primary tumours. Red: Methylated samples, Green: Unmethylated samples.

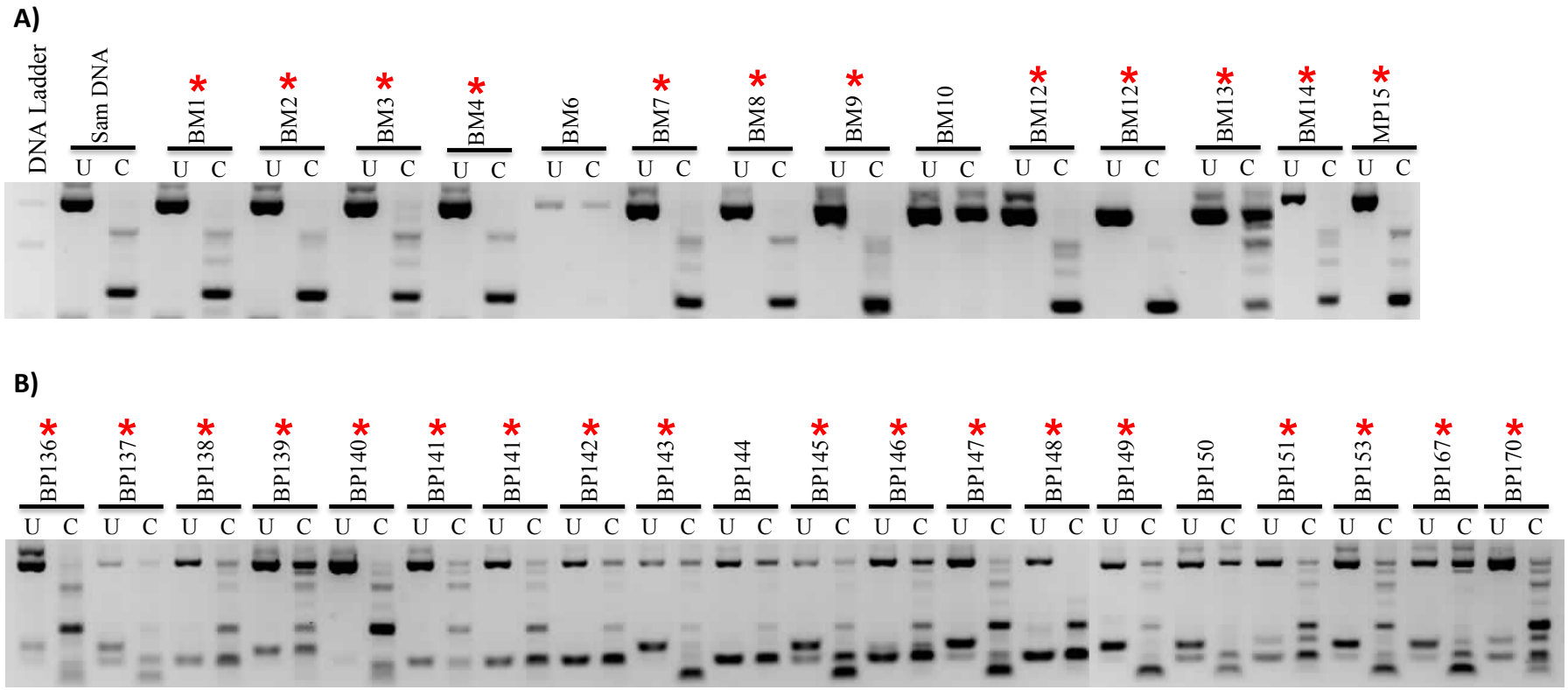


Figure 7.1: Methylation status of *HSPB9* in BBM (A), and in unrelated primary tumours (B). *HSPB9* is methylated in 12/14 (85%) in BBM samples and 2/17 (89%) *i.e.* it is methylated frequently both in BBM and primary breast tumours. Promoter region methylation was carried out using Combined Bisulphite and Restriction Analysis (CoBRA). Sam DNA: fully methylated positive control, BM: Brain metastases, BP: Primary breast tumours, U: Uncut/control sample, C: cut by restriction enzyme, *: methylated samples.

For the second approach, methylation status (β value) for each probe in all samples (normal breast tissues, primary breast tumours and the BBM) was generated. Comparison was carried out taking account of each probe in individual samples to determine the frequency of methylation in normal breast, primary breast tumours and the BBM sample. For each probe a β value of 0.4 was considered as the cut off value, so that a β value of >0.4 was considered as methylated whereas a β value <0.4 was considered as unmethylated (figure 7.2). This approach generated a list of 78 probes located in promoter regions of the genes that corresponded to 49 genes, of which, 27 probes (15 genes) were hypermethylated (increased level of methylation) in BBM samples compared to normal tissues and primary tumours whereas 56 probes (38) genes were hypomethylated (decreased level of methylation) in BBM samples compared to normal tissues and primary tumours (Figure 7.2). In addition to the list of probes located in promoter region of a gene, another list of differentially methylated probes were generated which also included probes both in promoter regions and in other regions of the gene such as gene body and 3'UTR in order to identify probes located other than in the promoter region of the genes (which could represent regulatory elements). This approach generated another list of probes, which are located in other genomic regions other than gene promoters. Not surprisingly, many of the probes were common in both the lists (as second list included probes in both promoter as well as other regions of a gene), therefore, only those probes which are not common with the probes in the promoter region and have a significant variations in methylation status (β value) were selected *i.e.* 7 probes (5 genes), of which 6 probes (4 genes) were hypermethylated whereas one probe (one gene) was hypomethylated. Therefore, this approach generated a long list of 85 probes (54 genes) in total, which are given in appendix E3. However, many of these probes (genes) did not have a significant difference in their methylation

status (β value) in BBM compared to primary tumours and normal tissues. Only 18 hypermethylated probes showed a significant difference (β value) in BBM compared to normal breast tissues and primary breast tissues, which were selected for experimental analyses (table 7.3, figure 7.3). Similarly, of 56 hypomethylated probes (38 genes), 16 probes (13 genes) showed significant variation in methylation status in normal tissues primary tumours compared to BBM samples. These probes were selected for laboratory analyses (Table 7.4 figure 7.4).

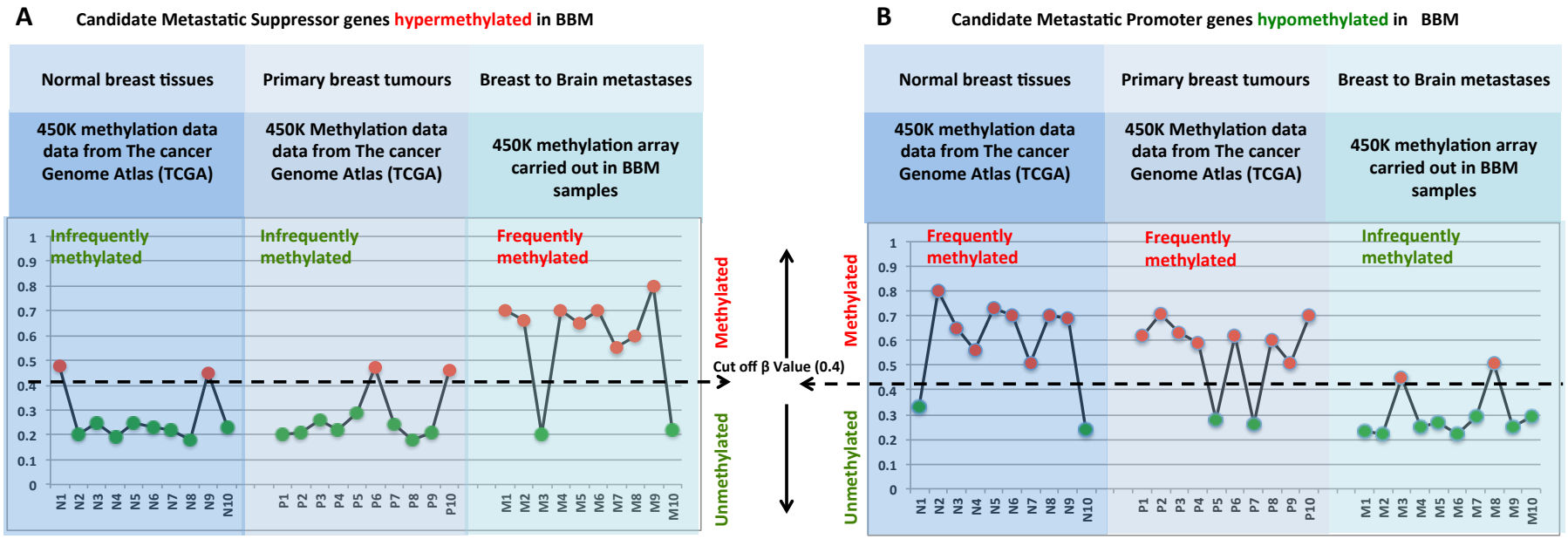


Figure 7.2: Analyses of the 450K-methylation array data identified two classes of genes dysregulated in BBM. A) Metastatic suppressor genes which are frequently methylated (β value ≥0.4 in ≥40% of the tumours) in BBM samples compared to normal breast tissues and primary breast tumours and B) Metastatic promoter genes which are infrequently methylated (β value < 0.4 in <40%) in BBM samples compared to normal breast tissues and primary breast tumours. The methylation data for the normal breast tissues and the primary breast tumours was downloaded from the TCGA where as the 450K-methylation array for BBM was carried out using 24 BBM sample.

Table 7.3: The list of probes selected for laboratory analyses, which were hypermethylated in BBM samples compared to normal breast tissues and primary breast tumours

Probes	Ensembl Gene ID	Gene	Chomosome
cg16736018	ENSG00000237588	*RP11-66D17.3*	*chr1*
cg01882471	ENSG00000239795	*AC109826.2*	*chr2*
cg15885430	ENSG00000250020	*RP11-811I15.1*	chr5
cg21532408	ENSG00000134864	*A2LD1*	chr13
cg12608565	*ENSG00000205710*	*C17orf107*	*chr17*
cg05529816			
cg14275842			
cg07834574			
cg09036188			
cg14482741			
cg20814095			
cg21071097			
cg08957069	ENSG00000221191	*AL662890.1*	*chr6*
cg04559779	ENSG00000207816	*MIR124−2*	chr8
cg20771240	ENSG00000254648	*RP11-713P17.4*	chr11
cg06769296			
cg22771759	ENSG00000234627	*NUS1P3*	*chr13*
cg17537073			

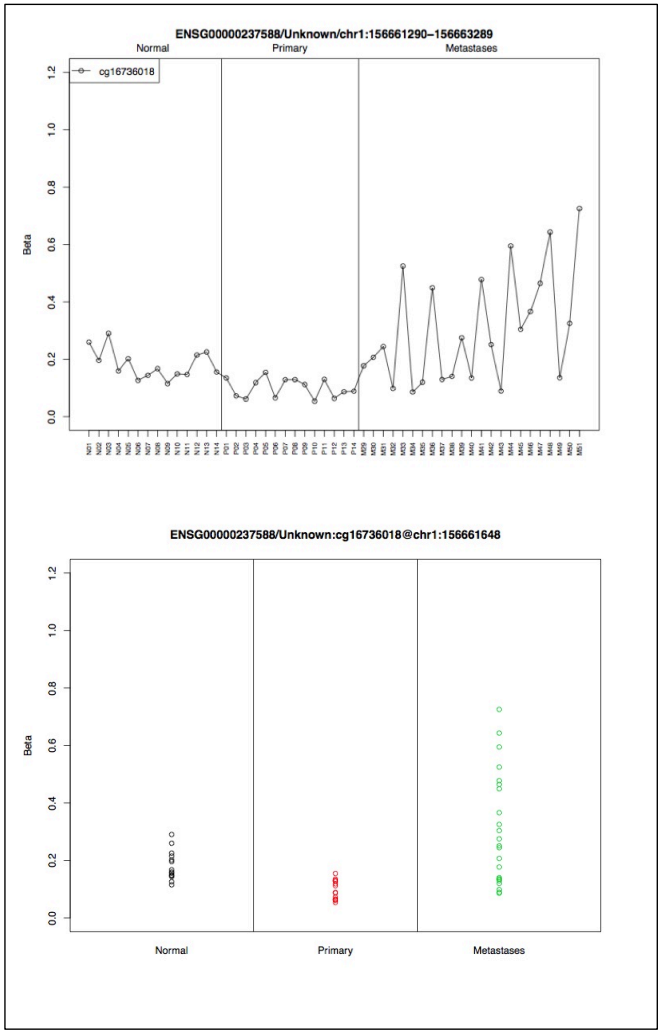
Table 7.3: The list of probers (genes) hypermethylated in BBM samples compared to normal breast tissues and primary breast tumours, which were selected for laboratory analyses for their methylation status. Methylation status of these genes was determined using CoBRA. Probes, which are located in regions other than the promoter of the gene, are shaded grey.

Table 7.4: The list of probes selected for laboratory analyses, which were hypomethylated in BBM samples compared to normal breast tissues and primary breast tumours

Probes	Ensembl Gene ID	Gene	Chomosome
cg25066665	ENSG00000163125	*RPRD2*	*chr1*
cg26362491	ENSG00000229367	*HMGN2P19*	*chr1*
cg00659878			
cg12720965	ENSG00000232192	*Unknown*	*chr1*
cg26563141	ENSG00000229604	*MT−ATP8*	*chr2*
cg13231117	ENSG00000229689	*AC009237.8*	*chr2*
cg27612889	ENSG00000138386	*NAB1*	*chr2*
cg12494166	*ENSG00000251129*	*RP11-734I18.1*	chr4
cg21771528	ENSG00000248693	*CTD-2023M8.1*	*chr5*
cg21806580			
cg23311108	ENSG00000185641	*Unknown*	*chr5*
cg12949141	ENSG00000249119	*MTND6P4*	*chr5*
cg24232869	ENSG00000226138	*SENP1/SUMO1*	*chr12*
cg09923107	ENSG00000264395	*MIR3193*	*chr20*
cg00494337			
cg23454038	ENSG00000247993	*FOXD1*	*chr5*

Table 7.4: The list of probers (genes) hypomethylated in BBM samples compared to normal breast tissues and primary breast tumours, which are selected for laboratory analyses for their methylation status. Methylation status of these genes was determined using CoBRA. Probes, which are located in regions other than the promoter of the gene, are shaded grey.

A cg16736018 (*RP11-66D17.3*)



B cg04559779 (*MIR124-2*)

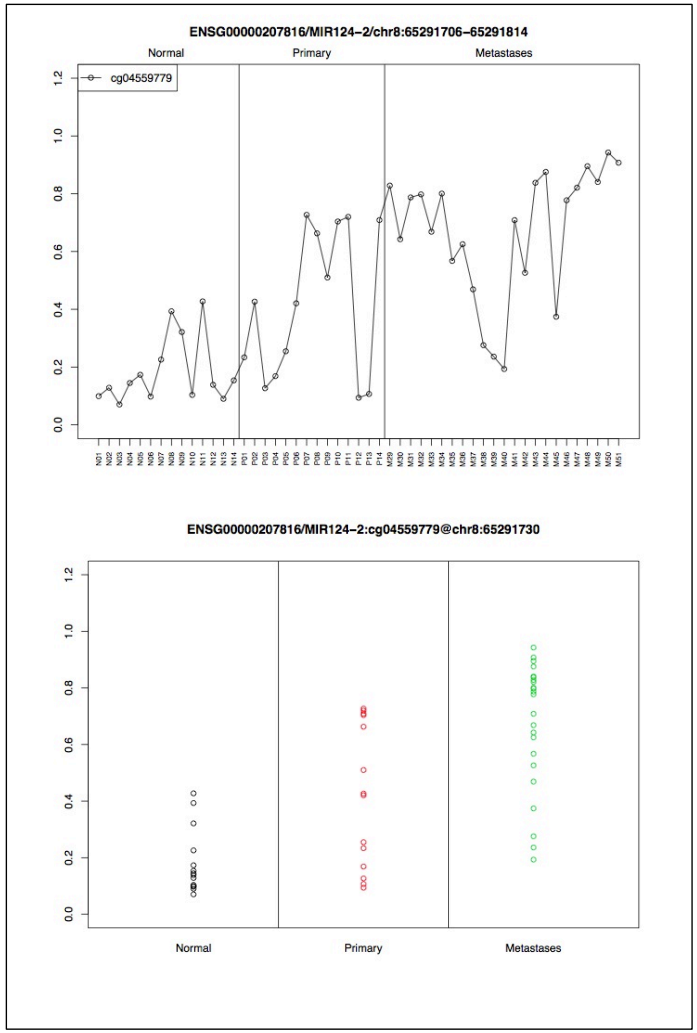
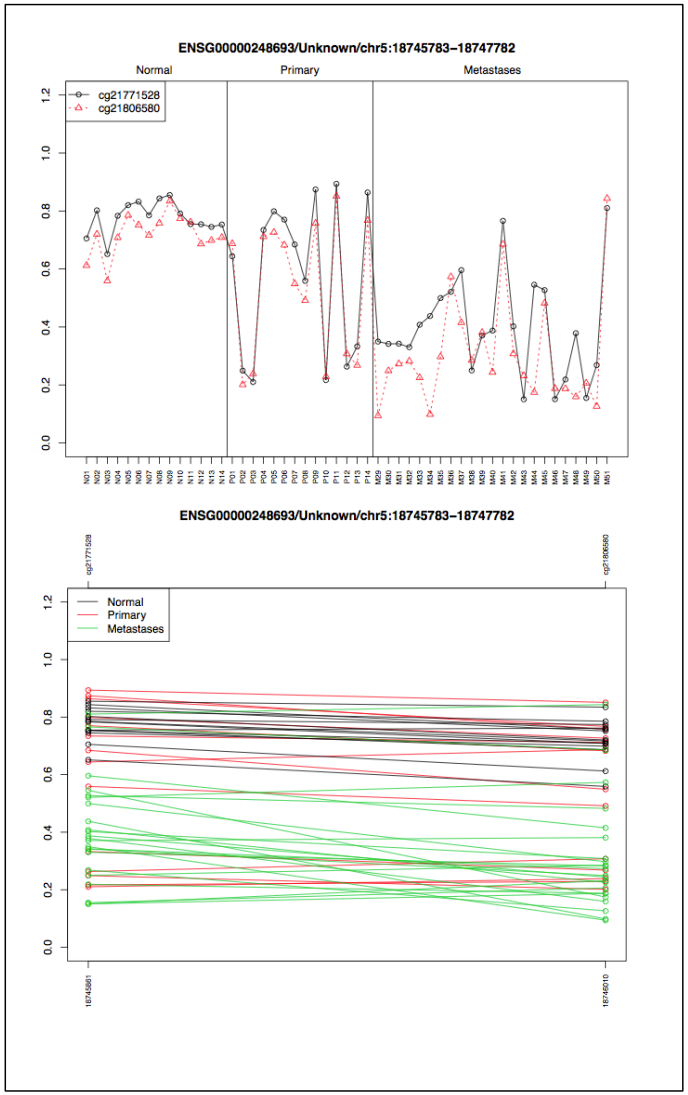


Figure 7.3: Representative examples of the probes, which are frequently methylated in BBM samples compared to normal breast tissues and primary breast tumours. This class of genes are the potential metastatic suppressor genes with increased level of methylation (hypermethylated) in BBM samples i.e. A) cg16736018 that corresponds to a non-protein coding gene RP1166D17.3, and B) cg207711240 and 06761296, both of which correspond to a non-protein coding gene RP11-713.4. The upper panel shows the β value of a probe in each individual sample whereas the lower panel shows the distribution of the probes based on the β value.

A cg21771528 cg21806580 (*CTD-2023M8.1*)



B cg12949141 (*MTND6P4*)

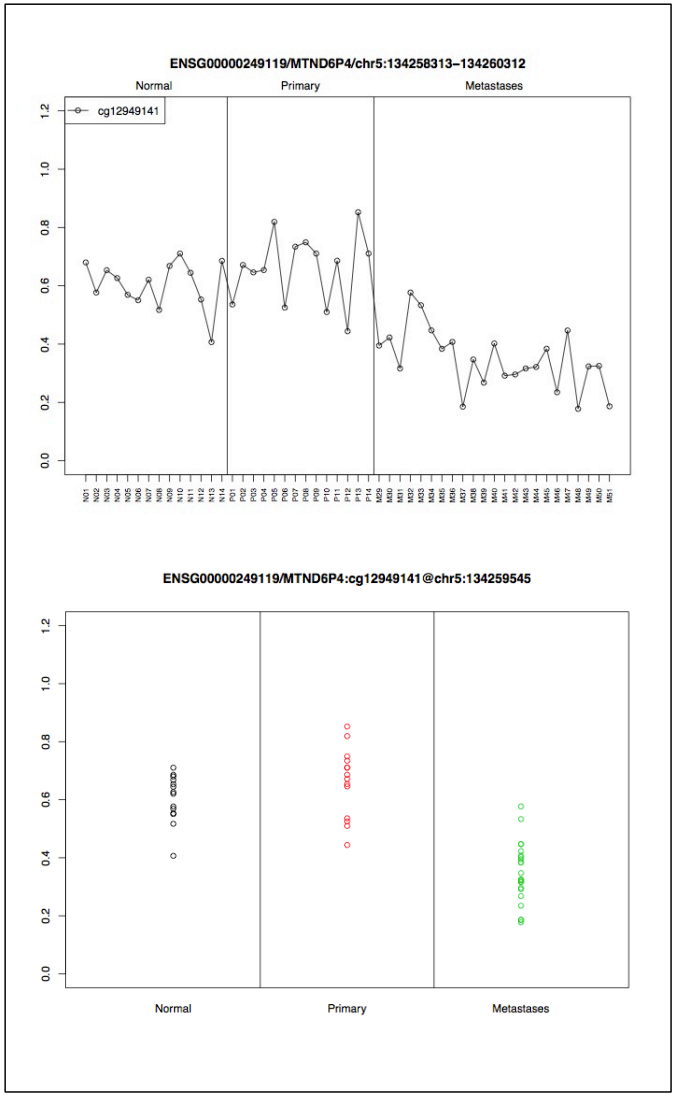


Figure 7.4: Representative examples of the probes, which are infrequently methylated in BBM samples compared to normal breast tissues and primary breast tumours. This class of genes are the potential metastatic promoter genes with decreased level of methylation (hypomethylated) in BBM samples *i.e.* A) cg21771528 and cg21806580 both of correspond to a non protein coding gene CTD-2023M8.1 and B) cg12949141 that corresponds to a pseudogene of MTND6 (MTND6P4) gene. The upper panel shows the β value of a probe in each individual sample whereas the lower panel shows the distribution of the probes based on the β value. The lower panel for gene 'A' looks different due to presence two probes in the same region.

7.2.3 Validation of candidate Hypomethylated probes in BBM samples

As an initial screening method, only those probe hypomethylated (infrequently methylated) in BBM were validated in laboratory using CoBRA. There were 16 probes (13 genes) that the methylation array analysis identified as significantly hypomethylated in BBM samples compared to normal breast tissues and primary breast tumours. These included one probe located in regions other than the promoter. A cohort of 15 BBM samples were used for initial validation to determine if these genes were infrequently methylated in the BBM samples analysed. From the panel of 13 genes analysed (including one probe hypermethylated), three genes (table 7.5) were infrequently methylated (≤40%) in BBM samples *i.e. MIR3193* (30%) and *CTD-2023M8.1 (20%) and FOXD1* (20%). Moreover, one gene *MTND6P4* (Figure 7.6) was not methylated in any of the BBM samples. These four genes were further validated in a cohort of 20 unrelated primary breast tumours (with no evidence of distant metastasis) to determine if they are frequently methylated (≥45%) (see section 7.2.4).

Table 7.5: Methylation status of candidate genes BBM

A)

Probe	Gene	Breast to Brain metastases (BBM) samples (n=15)															% Meth (n=15)
		BM1	BM2	BM3	BM4	BM5	BM6	BM7	BM8	BM9	BM10	BM11	BM12	BM13	BM14	BM15	
cg04559779	*mir124-2*																93
cg06769296 cg20771240	*RP11-713P17.4*																55
cg17537073	*NUS1P3*																40
cg09923107 cg00494337	*MIR3193*																33
cg21771528 cg21806580	*CTD-2023M8.1*																20
cg12949141	*MTND6P4*																0
cg23454038	*FOXD1*																33
cg12608565 cg05529816 cg14275842 cg07834574	*C17ORF107*																35
cg16736018	*RP11-66D17.3*																67
cg21532408	*AL2D1*																100
cg26362491 cg00659878	*HMGN2P19*																100
cg12494166	*RP11-734118.1*																100
cg15585430	*RP11-811115.1*																90
cg24232869	*SENP1/SUMO1*																100
cg08957069	*AL662890.1*																100

B)

Probe	Gene	Second cohort of BBM samples (n=16)															% Meth (n=31)
		BM16	BM17	BM18	BM19	BM20	BM21	BM22	BM23	BM24	BM25	BM26	BM27	BM28	BM29	BM30	BM31	
cg04559779	*mir124-2*																	88
cg06769296 cg20771240	*RP11-713P17.4*																	73
cg09923107 cg00494337	*MIR3193*																	30
cg21771528	*CTD-2023M8.1*																	26
cg12949141	*MTND6P4*																	0
cg23454038	*FOXD1*																	29

Table 7.5: Methylation status of array analysis candidate genes in BBM as determined by CoBRA. Of 14 genes analysed in an initial cohort of 15 BBM samples (A), seven genes (shaded in grey) were hypomethylated or hypermethylated (differentially methylated) in BBM compared to primary tumours. The methylation status of the other eight genes was the same in BBM and primary tumours. The seven differentially methylated genes were further investigated for their methylation status in a second cohort of 16 BBM samples (total of 31 BBM samples) (B). Red: methylated samples, Green: unmethylated samples, White: tumour samples did not show any product during CoBRA PCR.

7.2.4 Validation of candidate Hypermethylated of probes in a cohort of unrelated primary breast tumours

There were 18 probes (9 genes) hypermethylated in BBM compared to normal breast tissues and primary tumours, including 6 probes (4 genes) that were located in regions other than the promoter of the gene. These genes showed higher level of methylation (higher β value) in BBM samples, therefore, these genes were initially validated in an unrelated cohort of 20 primary breast tumours to investigate if they were infrequently methylated in primary tumours. Of 18 probes (9 genes), only three genes *RP11-713P17.4***,** *MIR124*-2, *NUS1P3* (Figure 7.5) were infrequently methylated in primary tumours. These genes were further validated in a cohort of 15 BBM samples (see section 7.3).

Table 7.6: Methylation status the candidate genes in a cohort of unrelated primary tumours

A)

Probe	*Gene*	Methylation Status of genes in a cohort of unrelated primary breast tumours (n=19)																			% Meth (n=19)
		BP136	BP137	BP138	BP139	BP140	BP141	BP142	BP143	BP144	BP145	BP146	BP147	BP148	BP149	BP150	BP151	BP153	BP167	BP170	
cg12608565 cg05529816 cg14275842 cg07834574	*C17ORF07*																				
cg26362491 cg00659878	*HMGN2P19*																				100
cg27612889	*NAB1*																				100
cg21532408	*AL2D1*																				100
cg12494166	*RP11-734I18.1*																				100

B)

		BP137	BP138	BP139	BP140	BP141	BP142	BP143	BP146	BP147	BP150	BP151	BP153	BP167	BP170	BP175	BP176	BP178	BP179	BP180	
cg04559779	*mir124-2*																				54
cg06769296 cg20771240	*RP11-713P17.4*																				11
cg17537073	*NUS1P3*																				7
cg23454038	*FOXD1*																				67

C)

		BP152	BP154	BP155	BP156	BP157	BP158	BP159	BP160	BP161	BP162	BP163	BP164	BP165	BP168	BP169	BP171	BP172	BP173	BP174	
cg09923107 cg00494337	*MIR3193*																				84
cg21771528	*CTD-2023M8.1*																				63
cg12949141	*MTND6P4*																				47

Table 7.6: Methylation status of candidate genes in a cohort of non-metastatic unrelated primary tumours to identify if they are frequently or infrequently methylated in these tumours. Seven genes (B, C) are differentially methylated in primary and BBM samples. *FOXD1* (B), *MIR3193*, *CTD-2023M8.1* and *MTND6P4* (C) are frequently methylated in primary tumours and are infrequently methylated in BBM samples (see table 7.5 A, C). Similarly, *NUS1P3* and *RP11-713P17.4* (B) are infrequently methylated primary tumours and are frequently methylated in BBM samples (see table 7.5A, C). *MIR124-2* (B) is frequently methylated both in primary tumours and in BBM (see table 7.5 A, C), however, the frequency of methylation of this gene is much higher in BBM samples (100%) compared to primary tumours (54%), which is statistically significant p=0.03). Five genes (A) are frequently methylated both in primary breast tumours and in BBM. Red: methylated samples, Green: unmethylated samples, White: tumour samples did not show any product during CoBRA PCR.

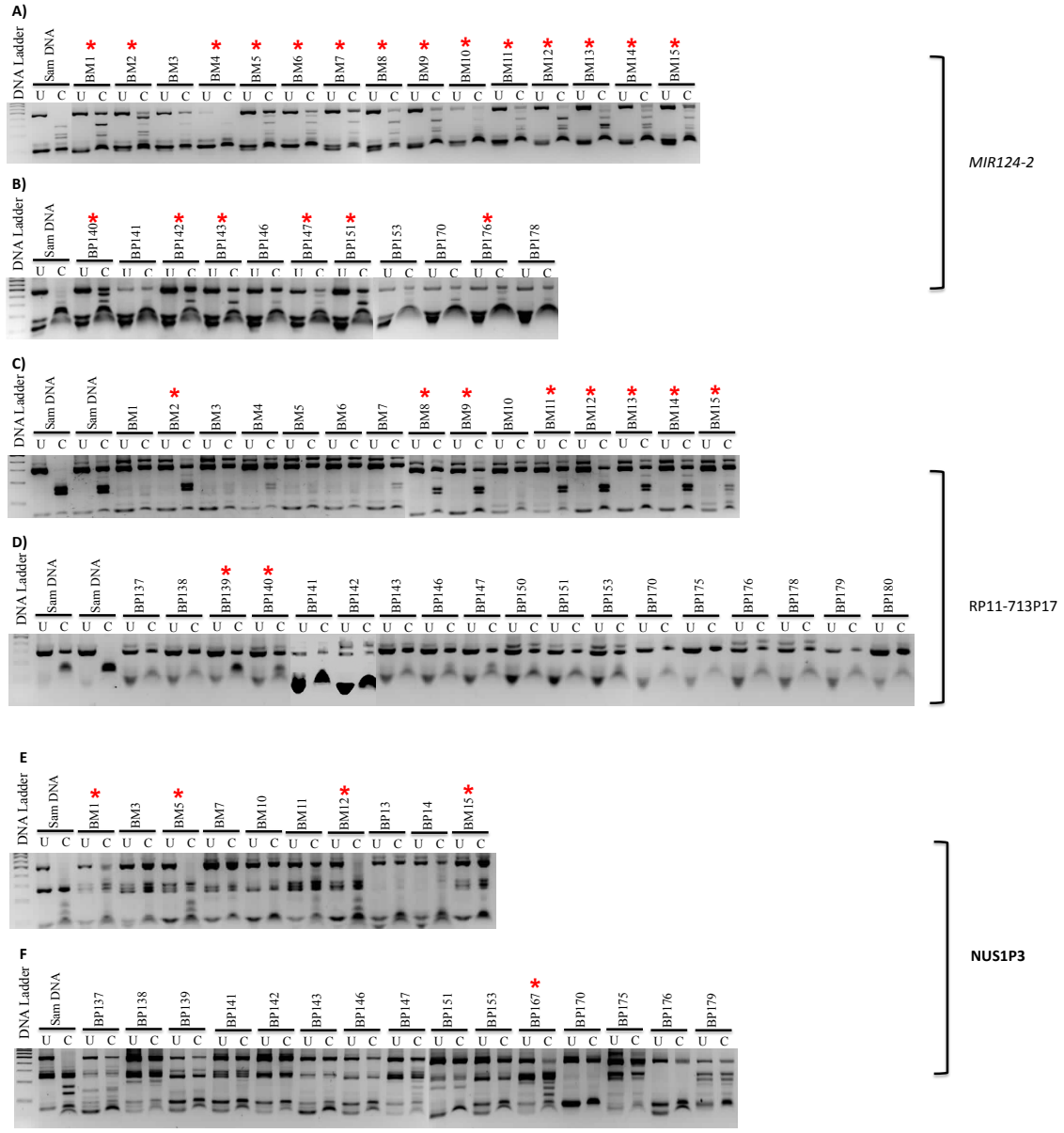


Figure 7.5: Methylation status of potential candidate metastatic suppressor genes in BBM versus unrelated primary tumours. *MIR124*-2 (A, B), *RP11*-713P17.4 (C, D) and *NUS1P3* (E, F) are enriched in methylation in BBM samples compared to unrelated primary breast tumours. *RP11-713P17.4* is frequently methylated in BBM (55%) and is infrequently methylated in a cohort of unrelated primary breast tumours (11%). *MIR12-2* is frequently methylated in BBM (93%) and primary breast tumours (55%), where as *NUS1P3* is infrequently methylated both in BBM samples (40%) and in primary breast tumours (7%) however, their methylation frequency is significantly higher in BBM compared to primary breast tumours (p=0.03 and 0.05 respectively). Sam DNA: Fully methylated positive

control, BM: Brain metastases, BP: Primary breast tumours, U: Uncut/control sample, C: cut by restriction enzyme, *: methylated samples.

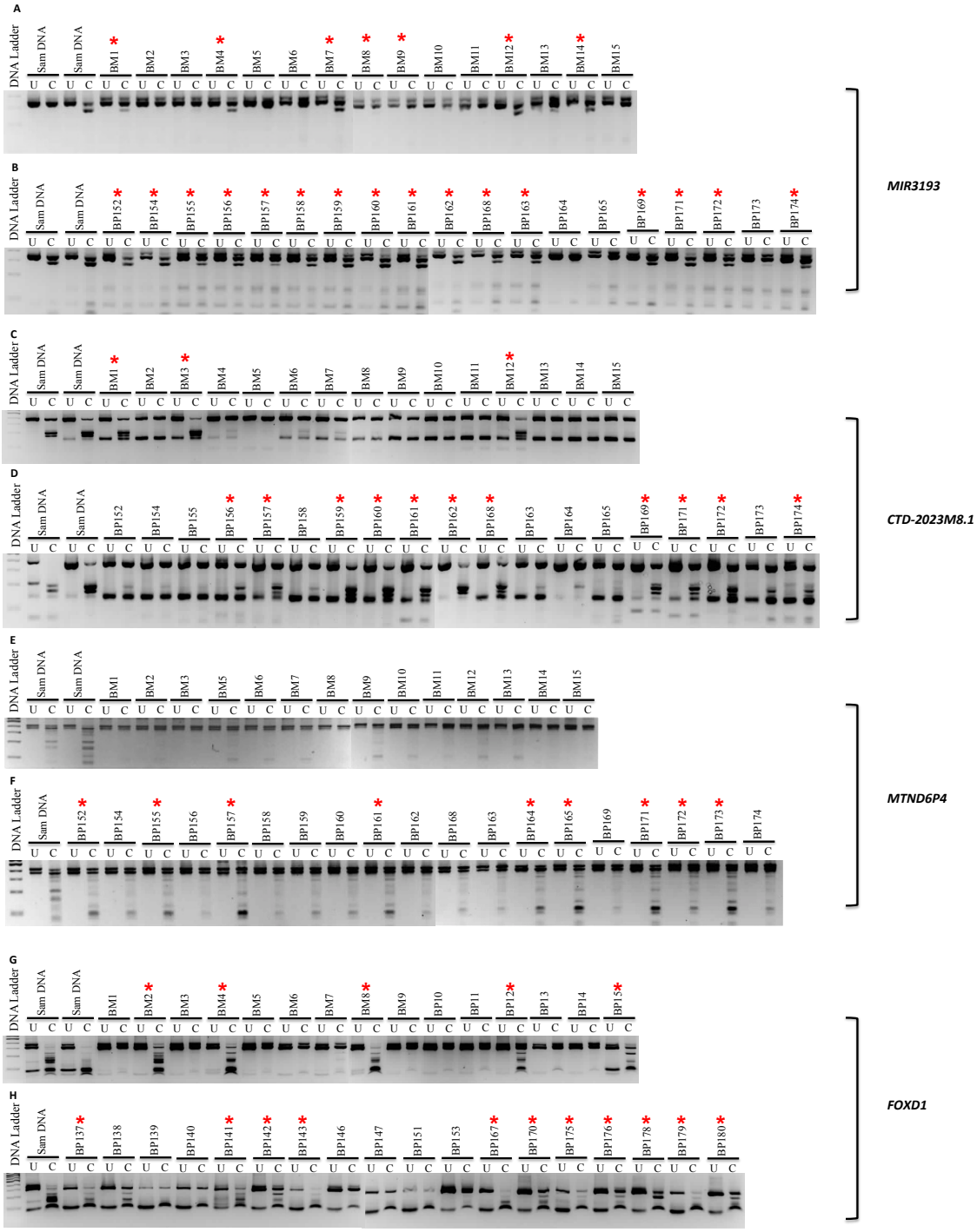


Figure 7.6: Methylation status of potential candidate metastatic promoter genes in BBM versus unrelated primary tumours. *MIR3193* (A, B), *CTD-2023M8.1* (C, D), *MTND6P4* (E, F) and *FOXD1* (G, H) have decreased methylation in BBM samples compared to unrelated primary breast tumours. *MIR3193* is frequently methylated in unrelated primary tumours (84%) and is infrequently methylated in BBM (30%). *CTD-2023M8.1* is frequently methylated in unrelated primary tumours (63%) and is infrequently methylated in BBM (20%). *MTND6P4* is frequently methylated in unrelated primary tumours (47%) where as it is not methylated in

any of the BBM samples analysed. *FOXD1* is frequently methylated in unrelated primary tumours (67%) and is infrequently methylated in BBM (33%). Sam DNA: fully methylated positive control, BM: Brain metastases, BP: Primary breast tumours, U: Uncut/control sample, C: cut by restriction enzyme, *: methylated samples.

Taken together, seven genes differentially methylated in BBM compared to primary breast tumours and normal breast tissues were identified (figure 7.7), of which three genes (*MIR124-2*, *RP11-713P17.4*, and *NUS1P3*) were hypermethylated in BBM (candidate metastatic suppressor genes) whereas four genes (*MIR3193*, *CTD-2023M8.1*, *MTND6P4* and *FOXD1*) were hypomethylated in BBM (metastatic promoter genes). The difference in the methylation of these genes in is statistically significant i.e. *MIR124*-2 (p=0.03), *RP11-713P17.4* (p=0.0001) (p, *NUS1P3* (p=0.05), *MIR3193* (p=0.01), *CTD-2023M8*.1 (p=0.0003), *MTND6P4* (p=0.01) and *MTND6P4* (p=0.0001). These genes were further investigated in BBM samples and their corresponding primary tumours in individual patients to investigate a possible point at which the epigenetic dysregulation occurred during the process of metastatic evolution.

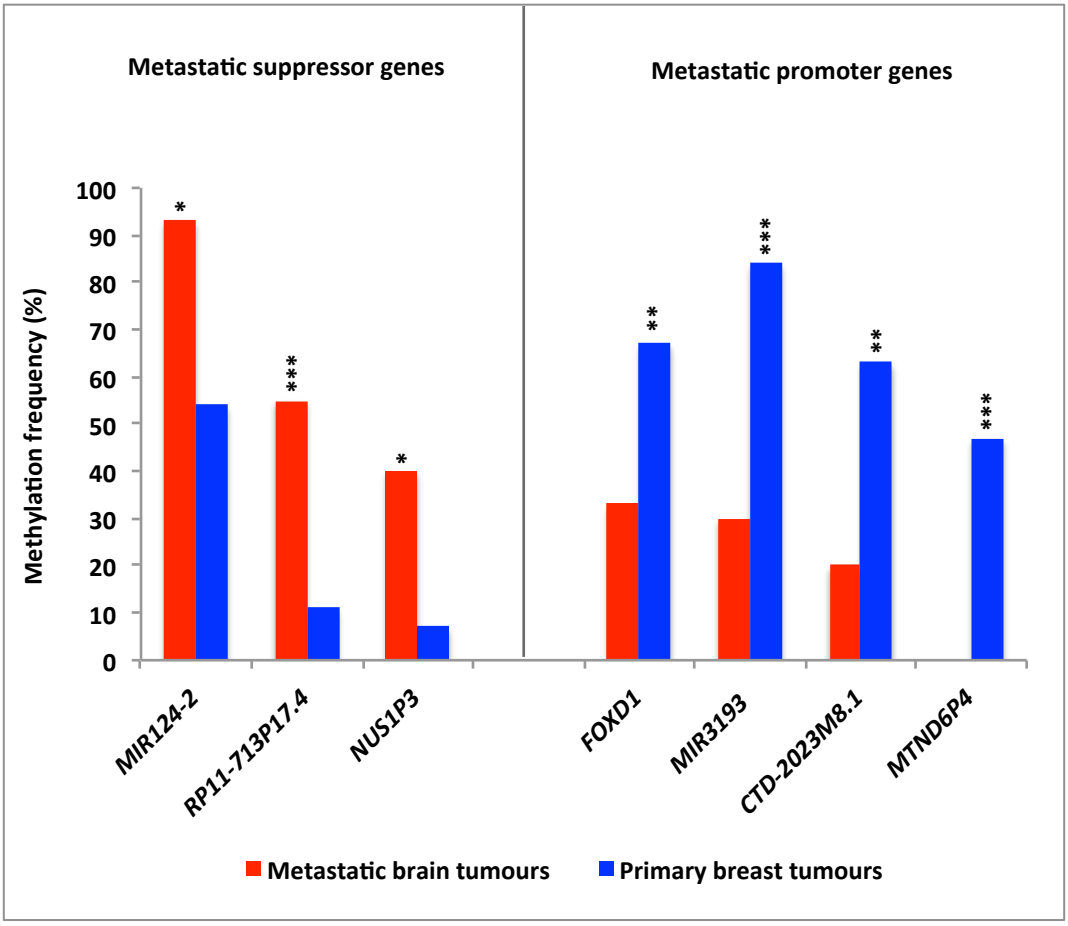


Figure 7.7: Methylation array analysis identified seven candidate genes differentially methylated in BBM compared to primary breast tumours and normal breast tissues, of which three were candidate metastatic suppressor genes hypermethylated and silenced in BBM samples whereas other four are candidate metastatic promoter genes hypomethylated and expressed in BBM samples. The difference in the methylation status of these genes in BBM compared to primary breast tumours is statistically significant (*: $p \leq 0.05$, **: $p \leq 0.01$, ***: $p \leq 0.001$).

7.2.5 Methylation status of the genes in BBM samples and their corresponding primary breast tumours from individual patients

The methylation status of the genes differentially methylated in BBM samples compared to normal breast tissues and the primary breast tumours were determined in matched pairs of tumours *i.e.* metastatic brain tumours and corresponding primary tumours from individual patients. Only 11 primary tumours were available from the 30 BBM patients, however, some loci in the primary tumour DNA proved refractive to amplification for some genes. 5 of 7 genes *i.e. MIR3193*, *FOXD1*, RP11-713P14.4, *CTD-2023M8.1* and *MTND6P4* were successfully amplified in BBM and their corresponding primary tumours from individual patients. In the matched samples *RP11-713P17.4* is methylated in 6 of 9 BBM and is commonly methylated in 4 of 9 primary breast tumours from these individual patients (Figure 7.7A). Of 9 successfully amplified matched pairs, *MIR3193* was methylated in only two of the brain metastases. However, it was methylated in 4 corresponding primary tumours (Figure 7.7B) Similarly, *FOXD1* was methylated in only 4 out of 11 matched BBM samples from the individual patients. However, it was methylated in all of the corresponding primary breast tumours (11/11) ((Figure 7.8A). *CTD-2023M8.1* was amplified in 5 matched pairs. It is methylated in 3 of 5 matched primary tumours; however, it is not methylated in any of the BBM originated from these primary tumours (Figure 7.8B). *MTND6P4* is amplified in 5 matched pairs, and is not methylated in any of the primary and BBM samples from individual patients (Figure 7.7C).

RP11-713P17.4 is methylated in 10% in unrelated primary tumours and 55% of the BBM. It is methylated in 67% (6/9) BBM samples and 50% (3/6) of the primary breast

tumours in which their corresponding BBM samples were methylated in individual patients (Figure 7.7A). This suggests that the methylation of *RP11-713P17.4* occurs early in the primary tumours before the tumours metastasise to the brain and the methylation is detectable in most primary tumours that metastasise to the brain, suggesting that it may play an important role in the early stages of primary tumour metastasis. *MIR3193* is methylated in only 3/9 (30%) of the primary breast tumours and 22% (2/9) in BBM in individual patients, where as it is methylated in 84% (16/19) in a cohort of unrelated primary tumours. This suggests that *MIR3193* methylation is decreased (hypomethylated) in BBM compared to primary tumours and the hypomethylation of *MIR3193* is detected in primary tumours from the individual patients further implying that the hypomethylation of *MIR3193* is an early event in BBM. Similarly, *MTND6P4* is methylated in 47% of unrelated primary tumours and is not methylated in any of the BBM samples. However, it is not methylated in any of the primary and BBM tumours in individual patients suggesting that the hypomethylated of *MTND6P4* occurs early in the primary tumours before the tumours metastasise to the brain.

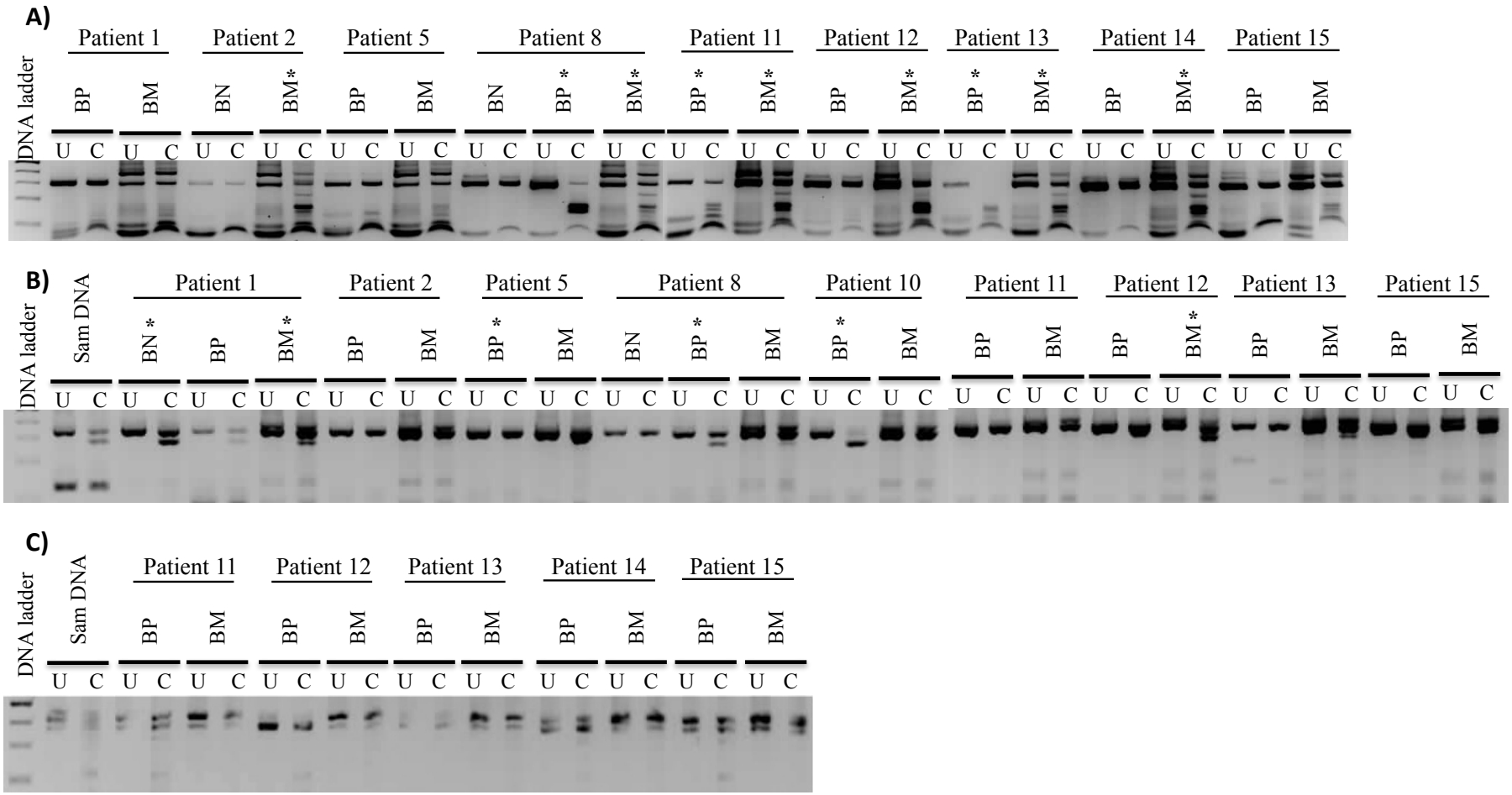


Figure 7.8: Methylation status (denoting early events) of *RP11-713P14.4 (A)*, *MIR3193 (B)* and *MTND6P4 (C)* in metastatic brain tumours and their corresponding originating primary breast tumours from individual patients. CoBRA was used to determine the methylation status; small, digested PCR products in the Bstu1 cut (C) lane compared to the undigested (U) lane indicates promoter

methylation in a sample. *RP11-713P14.4* is methylated in 6 BBM (see patient 2, 11, 12, 13, 14, 15) and all 3 corresponding primary tumours. One of the BBM (patient 2) has its only normal breast tissue amplified. Of 9 matched pairs analysed, *MIR3193* is methylated in only two of its BBM samples (patient 1 and 12) *i.e.* it is unmethylated in 7/9 BBM samples and is unmethylated in 6/9 corresponding primary tumours from individual patients. Interestingly, *MIR3193* is methylated in patient-1 in BBM and is not methylated in its matched primary tumour in contrast to all other samples. *MTND6P4* is not methylated in any of the BBM samples and their corresponding primary tumours from the individual patients. BP: Breast Primary tumour, BM: Metastatic Brain tumour, BN: adjacent Normal Breast tissue, U: Uncut/Control sample, C: cut by methylation specific restriction enzyme, *: Methylated samples, Sam DNA: fully methylated positive control.

In contrast, *FOXD1* is methylated in 67% of unrelated primary tumours and 28% of the BBM samples whereas it is methylated in all the primary tumours (11/11) and 36% of the BBM (4/11) from individual patients (Figure 7.8A). This suggests that hypomethylation of *FOXD1* occurs in the BBM at a late stage in the evolution of metastatic brain tumours, possibly after they have metastasised to the brain. Similarly, *CTD-2023M8.1* is methylated in 63% in unrelated primary tumours and 20% BBM samples. It is methylated in 60% (3/5) in primary tumours and none of the BBM from individual patients (Figure 7.8B). This suggests that the hypomethylation of these genes takes place at a late stage of tumour evolution in BBM. Alternatively, in these genes the genetic dysregulation such as hypomethylation of these genes may occur in a small subset of cells within the primary tumour (below the detection threshold of this assay) and these cells are enriched in the metastatic tumour. In summary, out of five genes that have been analysed in matched pairs in individual patients, dysregulation of three genes *MIR3193*, *MTND6P4* and *RP11-713P17.4* occurs early and is detectable in primary tumours (early events) where as the dysregulation of other two genes *FOXD1* and *CTD-2023M8.1* occurs late only after the primary tumour cells have left the primary tumours (late events).

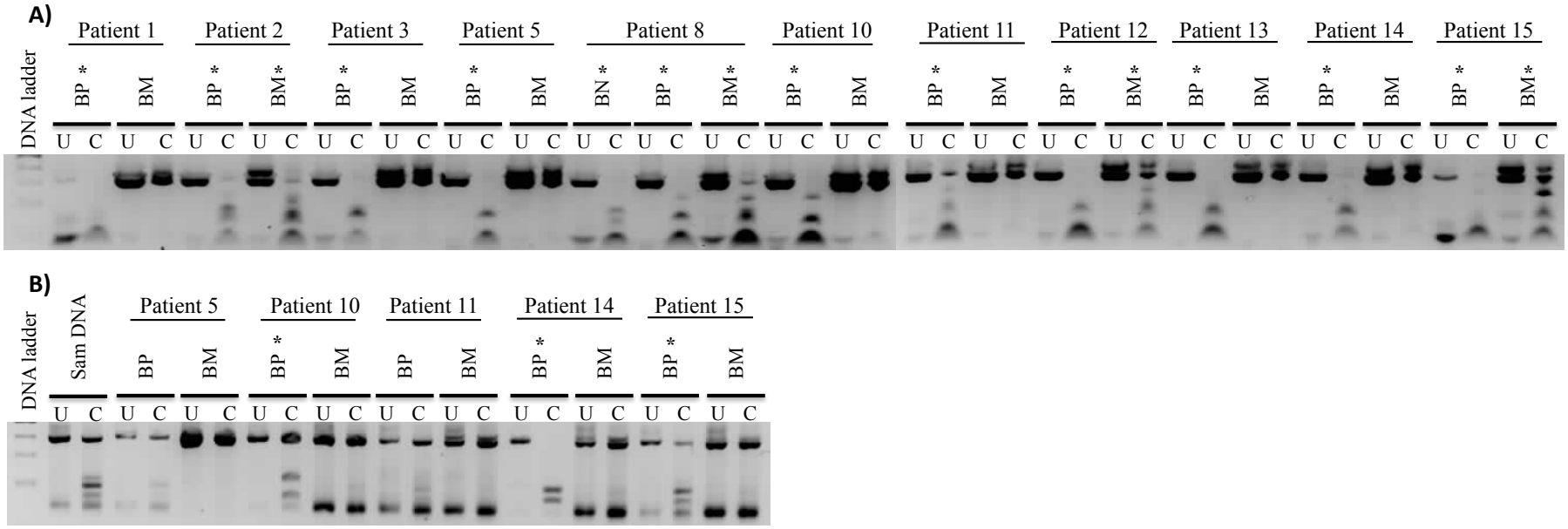


Figure 7.9: Methylation status (denoting late events) of *FOXD1 (A)* and *CTD-2023M8.1 (B)* in metastatic brain tumours and their corresponding originating primary breast tumours from individual patients. CoBRA was used to determine the methylation status; small, digested PCR products in the Bstu1 cut (C) lane compared to the undigested (U) lane indicates promoter methylation in a sample. *FOXD1* is methylated in all the primary tumours analysed where as it is methylated in only 4 of 11 of their corresponding BBM from the individual patients. Of 5 pairs analysed, *CTD-2023M8.1* is not methylated in any of the BBM samples where as it is methylated in their corresponding primary tumours from the individual patients. BP: Breast Primary tumour, BM: Metastatic Brain tumour, BN: adjacent Normal Breast tissue, U: Uncut/Control sample, C: cut by methylation specific restriction enzyme, *: Methylated samples, Sam DNA: fully methylated positive control.

7.2.6 Investigating the methylation status of genes in patients' plasma to determine if these genes could be used as prognostic markers for BBM

The 450K-methylation array identified seven genes; *MIR3193*, *FOXD1*, *MTND6P4*, *CTD-2023M8.1*, *RP11-713P17.4*, *MIR124*-2 and *NUS1P3* that were differentially methylated in BBM compared to primary tumours and normal breast tissues. Patients' plasma, collected at the time of BBM surgery was used to investigate if these genes could be used as non-invasive prognostic markers. To investigate this, tumour-free circulating DNA was extracted from 10 plasma samples to carry out methylation specific PCR (MSP) in six of these genes (analysis of *NUS1P3* was not carried out due to time constraints) (Table 7.7). In addition, MSP was also carried out for *CCDC8*, which was identified as a candidate by literature review (section 4). It was expected that the methylation status of the BBM to be similar to that of the serum as it is possible that the tumour-free circulating DNA in the patients' plasma is the results of the DNA sloughed off from the BBM (Esteller *et al.*, 1999; Hoque *et al.*, 2006; Lo Nigro *et al.*, 2013) due to tumour-associated leaky blood brain barrier (Zhang & Yu, 2011). To ensure the uniformity in the methodology used, MSP was carried out on BBM samples corresponding to those plasma samples from the individual patients.

CCDC8 was methylated in 9/10 (90%) brain metastases according to MSP in BBM samples. This gene is methylated in 6/9 (67%) plasma samples in these patients (Figure 7.9A). There is a variation in methylation status of *CCDC8* between BBM and plasma samples in three patients (BM8, BM11 and BM13). Methylation status of *MIR3193* in 50% of BBM is identical to the methylation status in serum samples (Figure 7.9B). Interestingly, methylation of *MIR3193* was detected more frequently in plasma samples

(50%) compared to BBM samples. Similarly, *CTD-2023M8.1* is methylated in 6/10 BBM samples. It is methylated in those plasma samples from the patients whose BBM were methylated. This showed that the methylation status of *CTD-2023M8.1* in BBM samples is identical to the methylation status of the free circulating DNA in plasma samples in individual patients (Figure 7.9C). *RP11-713P17* was methylated in 8/10 plasma samples with a variation of only one sample between BBM and plasma samples. *MIR124-2* is methylated in all the BBM samples, and it is methylated in 8/10 plasma samples among this patients (Figure 79D). There were two patients BM12 and BM13 that showed variations in methylation status *of MIR124-2* in BBM and plasma samples.

It will be necessary to carry out MSP of circulating DNA for *GALNT9, BNC1 and MTND6P4* to get a complete picture of the detectability of tumour methylation in all candidate genes in BBM and serum DNA.

This analysis has shown that there is a good correlation between the methylation status of circulating DNA in plasma and fresh frozen BBM tumours in individual patients. This suggests that it might be possibility to use these genes as a panel of potential biomarkers for BBM with suitable quantitative methylation techniques such as QMSP. This analysis has furthermore identified a range of novel candidate prognostic markers. MIR124-2 (a tumour suppressor microRNA associated with various malignancies), RP11-713P17.4 (a novel noncoding RNA gene; ncRNA) and NUS1P3 (a pseudogene of its parental pro-apoptotic gene; NUS1) are potential metastatic suppressor genes. Similarly, MTN6P4 (a processed pseudogene of its parental gene mitochondrial encoded NADH dehydrogenase; MTND6), *CTD-2023M8.1* (a novel non protein gene/noncoding RNA gene), *MIR3193* (a novel microRNA identified by deep

sequencing in melanoma) and *FOXD1* (a member of forkhead transcriptional factors) are potential metastatic suppressor genes. Furthermore, three novel metastatic suppressor genes *BNC1* (a transcriptional factor), *CCDC8* (a regulator of microtubule dynamics) and *GALNT9* (initiator of o-glycosylation) are potential metastatic suppressor genes see section 4.3 and 5.3 for detailed function of these genes).

Table 7.7: Methylation status of candidate prognostic markers in BBM and tumour free circulating DNA

Patient	*CCDC8*		*MIR3193*		*CTD*		*FOXD1*		*RP11*		*MIR124-2*		Methylation status mets vs serum (% identity)
	MSP Mets	MSP ser	MSP Mets	MSP ser	MSP Mets	MSP ser	MSP Mets	MSP ser	MSP Mets	MSP ser	MSP Mets	MSP ser	
1													100
2													100
5													100
6													67
7													75
8													83
10													83
11													50
12													75
13													50

Table 7.7: Methylation status of candidate prognostic markers in BBM and serum DNA determined by Methylation Specific PCR (MSP). Methylation status of panel of 6 candidate prognostic markers in individual patients in their BBM samples versus serum DNA has been determined. It is expected that the methylation status of each gene in BBM should be identical with the serum in each patient. The identity (BBM vs. serum) in methylation has been expressed in percentage for each patient in a panel of 6 candidate prognostic markers.

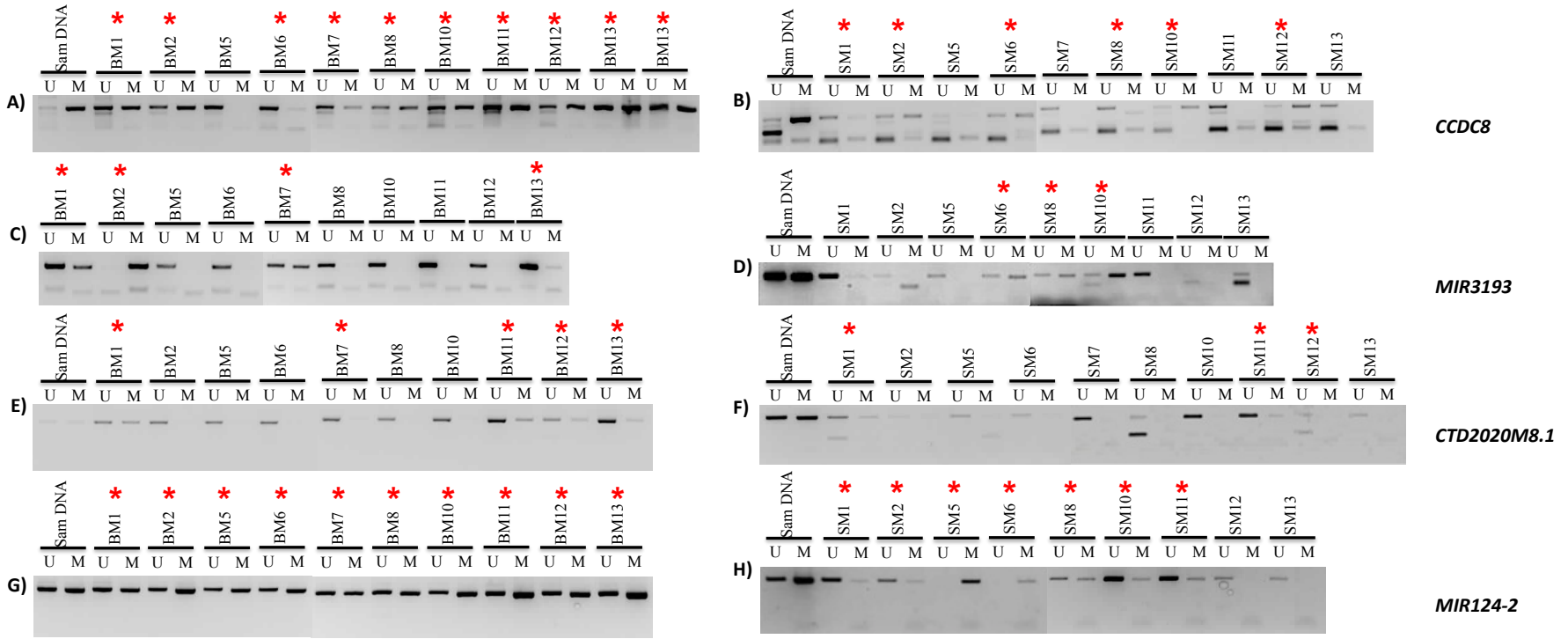


Figure 7.10: Methylation status of candidate prognostic markers (genes) CCDC8 (A, B), MIR3193 (C, D), CTD-2023M8.1 (E, F) and MIR124-2 (G, H), in BBM versus tumour free serum DNA. Methylation status of BBM is similar to serum DNA in individual patients in majority of the samples suggesting that serum the methylation status of serum DNA could possible used as prognostic markers for BBM. BM: Brain Metastases, SM: Serum DNA U: USP, M: MSP, *: Methylation

7.3 Discussion

In recent years, a wide range of epigenetic studies, especially those concerning the DNA methylation status of primary breast tumours have been carried out (Yuan *et al.*, 2003; Huang *et al.*, 2011; Moelans *et al.*, 2011; Park *et al.*, 2011; Zeng *et al.*, 2012; Kar *et al.*, 2014; Shenker *et al.*, 2015). However, due to the poor clinical outcomes of breast cancer patients, there is an urgent need to understand the process of BBM and to identify a way of prognosis or to find novel targeted therapies (Salhia *et al.*, 2014). Due to the challenges associated with the treatment of brain tumours such as drug access via the blood brain barrier and other related problems (Fazakas *et al.*, 2011), it is crucial to find novel prognostic markers that will inform the clinical management of breast cancer patients.

There are very few prognostic procedures available for BBM; however, these are primarily based on the various factors including ethnicity, ER/PR/HER receptor status and presence of germline mutations *BRCA1* gene (Salhia *et al.*, 2014). However, there are not many genome-wide studies carried out to find rare epigenetic determinants contributing to BBM. Salhiya *et al* (Salhia *et al.*, 2014) carried out copy number analyses, mRNA expression profiling and 27K methylation studies to find dysregulated gene in BBM which however, did not identify any potential prognostic biomarkers or key regulators that were significantly differentially dysregulated in BBM compared to primary tumours and normal breast tissues.

This study aims to find dysregulated genes, by genome-wide analyses that are involved in the process of BBM. For this purpose, Illumina BeadChip 450K human methylation arrays were carried out using 24 fresh frozen samples from patients with BBM. In

addition, genome wide 450k methylation data for primary breast tumours and normal breast tissues was downloaded from the TCGA to identify differentially methylated probes/genes in BBM compared to normal tissues and primary tumours.

This methylation array analyses generated a long list of 78 probes located in promoter regions as well as 7 probes located other than the promoter regions. Of 85 probes in total, only 34 probes (21 genes), showed clear differences in methylation in BBM compared to primary tumours and normal breast tissues. Experimental validation of methylation of these genes identified only 7 genes (*MIR3193*, *FOXD1*, *MTND6P4*, *CTD-2023M8.1*, *RP11-713P17.4*, *MIR124*-2 and *NUS1P3*) that were differentially methylated in BBM compared to primary tumours and normal breast tissues. Of these 7 genes, *RP11-713P17.4*, *MIR124*-2 and *NUS1P3* were frequently methylated in BBM, and infrequently methylated in primary breast tumours and normal breast tissues where as *MIR3193*, *FOXD1*, *MTND6P4* and *CTD-2023M8.1* were frequently methylated in normal breast tissues and primary tumours, and infrequently methylated in BBM.

One of the important findings of this study is the identification of novel regulatory elements and microRNAs dysregulated in BBM. Of 7 genes identified, there were two microRNAs (*MIR3193* and *MIR124-2*), four non-coding genes that included two long intergenic non-coding RNA (lincRNA) genes (*RP11-713P17.4* and *CTD-2023M8.1*) and two non-protein coding genes or the pseudogenes (*MTND6P4* and *NUS1P3*).

FOXD1 codes for a protein forkhead box protein D1, a member of a family of forkhead transcriptional factors containing distinct forkhead domain (Ernstsson *et al.*, 1996; Lancet *et al.*, 2013). *FOXD1* is identified as a mediator and indicator of gene expression changes during the cell reprogramming process (Koga *et al.*, 2014). The population of induced pluoripotent stem cells (iPSCs) is reduced after *FOXD1* is knockdown. *DAX1*,

which is a component in autoregulatory network of maintenance of pluripotency in the cells is inhibited by knockdown of *FOXD1*, and furthermore, the iPSCs colonies are a result of *FOXD1* expression, revealing that *FOXD1* contributes to iPSCs populations for cell programming process (Koga *et al.*, 2014). It has been shown that the forkhead genes are associated with nervous system development, tumorigenesis as well as in gene expression in liver and lung tissues (Ernstsson *et al.*, 1996). It is expressed in temporal regions of the developing chick retina, determining the temporal specificity of retina (Takahashi *et al.*, 2009; Carreres *et al.*, 2011). Moreover, *FOXD1* is known to act as a transcriptional repressor of downstream gene in retinal development (Takahashi *et al.*, 2009). Interestingly, co-transfection of *FOXD1* with *p53* has revealed that, in presence of p53, *FOXD1* gene inhibits the transcription of genes, which otherwise would be activated by it in absence of p53 suggesting that *FOXD1* regulation is associated with p53 (Ernstsson *et al.*, 1996). In addition, co-transfection of *WT-1* with *p53* has produced similar results showing that *WT-1* is regulated by *p53* and a binding site for *WT-1* is found in in *FOXD1*. These experiments have demonstrated that the two genes *p53* and *WT-1* regulate the *FOXD1* gene (Ernstsson *et al.*, 1996). These data are consistent with our findings that *FOXD1* is a metastatic promoter gene silenced in normal breast tissues and primary breast tumours through promoter methylation. In addition, the two tumour suppressor genes *WT-1* and *p53* may provide an additional layer of control on this gene. *FOXD1* is unmethylated in BBM and is methylated in primary tumours in the individual patients suggesting that the hypomethylation and expression of *FOXD1* gene takes place only after the tumours have metastasised to the brain. *FOXD1* could possibly contribute to the tumour survival in the brain by promoting tumour cell proliferation through pluoripotent stem cells (PSC) or by stimulating *DAX1*, which helps to maintain PSC and helps in cell reprogramming.

MIR124-2 is one of the *MIR124* precursor gene located in chromosome 8 (Deo *et al.*, 2006; Wang *et al.*, 2014b). There are three independent precursors of *MIRNA124* gene *i.e. MIR124-1*, *MIR124-2*, and *MIR124-3* located on chromosome 8, 8 and 20 respectively (Conaco *et al.*, 2006; Deo *et al.*, 2006; Kozomara & Griffiths-Jones, 2014; Neo *et al.*, 2014; Wang *et al.*, 2014b), which are processed to form *MIR124*. *MIR124* is one of the most abundantly expressed and conserved microRNAs in the nervous system (Deo *et al.*, 2006; Makeyev *et al.*, 2007). Interestingly, expression of *mir124* is specific to neurons and its expression level increases as nervous system develops. Moreover, neuronal cell differentiation is activated where as glial cell differentiation is repressed by *MIR124*. *MIR124* promotes neurogenesis by down-regulating *SCP1* (small C-terminal domain phosphatase 1), which has an anti-neural function during neural development (Visvanathan *et al.*, 2007). Therefore, *MIR124* is involved in regulation of intricate nervous system network specific alternative splicing and plays a crucial role in progenitor cells differentiation into the mature neurons (Deo *et al.*, 2006; Makeyev *et al.*, 2007). *MIR124*, furthermore, down regulates non-neuronal mRNA when introduced into neuronal cells, suggesting its specific role in promoting neuronal and inhibiting role in non-neuronal cell differentiation (Makeyev *et al.*, 2007). *MIR124-2*, one of the genomic loci of *MIR124* is downregulated due to promoter hypermethylation in hepatocellular carcinoma (HCC) cell lines (Furuta *et al.*, 2010). In addition, restoration of *MIR124* in two different HCC cell lines had suppressive effects on cell proliferations suggesting that *MIR124* supresses cell proliferations. In addition, its expression led to decrease in S, G2 and M cells and increase in G0-G1 cells suggesting that *MIR124* contributes to cell cycle arrest at G1-S phase checkpoint possibly due to apoptosis of HCC cells. (Furuta *et al.*, 2010). Previous studies (Han *et al.*, 2013) identified that *MIRNA124* directly targets the 3'UTR region and supresses *CD151*, which is highly

expressed in breast cancers and is associated with TGF-β signalling. *CD151* promotes motility of breast cancer cell lines contributing to breast cancer migration and invasion (Han *et al.*, 2013). Similarly, mimic of *MIR124* and knockdown of *CD151* together leads to the inhibition of breast cancer cell lines through cell cycle arrest (Han *et al.*, 2013) where as *MIR124* mimics alone increases breast cancer motility. Metastases of breast cancer cell lines to lung are diminished by knockdown of *CD151* and its transfection increases the migration and invasion of breast cancer cell lines. In addition, motility of breast cancer cell lines is inhibited by overexpression of *MIR124* suggesting that *MIR124* regulates breast cancer metastases and invasion by targeting *CD151* (Han *et al.*, 2013). Similarly, *MIR124* inhibits motility and adhesion properties of oral squamous cell carcinoma (OSCC) by targeting *ITGB1*, which is a crucial factor for OSCC progression and metastases (Hunt *et al.*, 2011). Furthermore, reduced expression of *MIR124* is associated with poor prognosis of colorectal carcinoma (CRC) and poor outcome of the patients (Wang *et al.*, 2013b; Jinushi *et al.*, 2014). *In vitro* and *in vivo* knockdown experiments have shown that *MIR124* is expressed in microglia as well as the neuronal population of the normal CNS, regulates microglia quiescence restricting their proliferation (Ponomarev *et al.*, 2011; Willemen, 2012) (Conaco *et al.*, 2006; Deo *et al.*, 2006; Akerblom *et al.*, 2012) and supresses macrophages and monocytes differentiation into microglia, which suggests that *MIR124* may be a master regulator of different cell types in CNS development (Ponomarev *et al.*, 2011). All three *MIR124* genes; *MIR124-1*, *MIR124-2* and *MIR124-3* are hypermethylated and silenced in pancreatic adenocarcinoma (PDAC) and are associated with migration and invasion of PDAC (Wang *et al.*, 2014b). *MIR124* is associated with downregulation of its target gene *RAC1* that is involved in adhesion, migration and invasion of PDAC (Wang *et al.*, 2014b). Similarly, stable decrease in *MIR124* is shown to promote migratory and

invasive behaviour of the pulmonary fibroblast cells suggesting the suppressive role of *MIR124* in migration and invasion of cancer cells (Wang *et al.*, 2014a). Taken together, *MIR124*, which is processed from any of its three genes (*MIR124-1*, *MIR124-2* and *MIR124-3*) is abundantly expressed in neuronal cells and glial cells contributing to glial cells quiescence and is involved repression of migration and invasion of various cancers through its target. These findings are consistent with our finding that silencing of *MIR124-2* through promoter methylation in BBM samples may provide a selective advantage for metastasised tumours to survive and to proliferate in the brain. In addition, it's silencing in primary breast tumours may contribute to its metastases from the breast to brain and other sites in the body. Moreover, it could be possible that *MIR124-2* silencing in BBM may provide a selective advantage to the tumours to proliferate in non-neuronal cellular environment such as amongst glial cells. It is important to note that methylation of *MIR124-1* was higher in BBM compared to primary tumours; however, it was infrequently methylated both in BBM and primary breast tumours (section 4.2). This suggests that among the members of *MIR124* genes (*MIR124-1*, *MIR12-2* and *MIR124-3*), only *MIR124-2* is significantly dysregulated in BBM due to promoter hypermethylation.

MIR3193 is a novel microRNA molecule, which is one of the 209 micro RNAs identified by deep sequencing melanomas (Stark *et al.*, 2010). There is no other evidence of *MIR3193* dysregulation reported in any cancer types or other diseases. In this study, *MIR3193* was frequently methylated in non-metastatic primary breast tumours and infrequently methylated in BBM samples. In addition, it is commonly unmethylated in primary breast tumours and BBM samples in individual patients suggesting that the demethylation/hypomethylation of *MIR3193* is an early event during

the course of tumour evolution. It implies *MIR3193* is a metastatic promoter microRNA, which is kept silenced in normal breast tissues or in primary breast tumours through the methylation. *MIR3193* is hypomethylated in those subsets of cells or tumours, which metastasises to the brain. The hypomethylation of *MIR3193* and its expression in primary breast tumours is an early event, which could possibly contribute to the process of BBM.

MTND6P4 and *NUS1P3* are non-coding genes or processed pseudogenes of its parental gene *MTND6* and *NUS1* respectively. *MTND6*, which codes for the protein mitochondrial encoded NADH dehydrogenase, provides a quinone binding sites of complex I in mitochondria in ETS (DeHaan *et al.*, 2004) and is one of the six subunits (*ND1-ND6*) of the complex I in mitochondria (Lancet *et al.*, 2013). Furthermore, activating mutations in *MTND6* are associated with increase in metastatic potential in lung and breast cancer cell lines (Ishikawa *et al.*, 2008). Similarly, *NUS1*, also known as *NOGO-B*, an isoform of *NOGO* gene is expressed in central and peripheral nervous system including lung, kidney, heart and liver as well as in other tissues (Huber, 2002; Schwab, 2010). In contrast to the previous studies which showed that *NUS1* is a pro-apoptotic gene (Li *et al.*, 2001), it has been shown that the *NUS1* down regulates epithelial markers such as E-cadherin and increases mesenchymal markers contributing to EMT in cervical cancer promoting invasion and metastasis (Xiao *et al.*, 2013). In addition, *NUS1* has been associated with cell cycle progression (Harrison *et al.*, 2011) and its expression has been associated with various cancer types (Oertle *et al.*, 2003). Similarly, SUN1 expression is associated with proliferation of ER/PR/HER2 positive breast tumours (Wang *et al.*, 2013a).

Previous studies (Poliseno *et al.*, 2010) of the *PTEN* pseudogene, *PTENP1* has revealed

that the *PTENP1* has highly identical sequence with its ancestral gene *PTEN*, due to which both the *PTEN* and *PTENP1* compete for their miRNA binding for their decoy in the cells (Poliseno *et al.*, 2010). Therefore, increase in the expression of *PTENP1* led to the increase in the *PTEN* due to increased availability of *PTENP1* for miRNA binding (Poliseno *et al.*, 2010). It is plausible that such regulation is common and the genes identified in this study mediate the regulation of their partner genes this way.

MTND6P4 is frequently methylated in unrelated primary tumours and infrequently methylated in BBM samples. In addition, *MTND6P4* is commonly unmethylated in BBM patients and their corresponding primary tumours from the individual patients. This suggests that *MTND6P4* is switched off in the normal breast tissues due to methylation and is turned on due to hypomethylation in primary breast tumours. The hypomethylation of *MTND6P4* and its expression in primary breast tumours is an early event, which could possibly contribute to the process of BBM. This is consistent with the idea that increase in the expression of *MTND6P4* leads to the increase in expression *MTND6* transcripts and proteins that in turn increases the metastatic potential of the primary breast cancer cells contributing to BBM. Similarly, *NUS1P3* is infrequently methylated in primary tumours suggesting its expression, which possibly could lead to increased expression of *NUS1*. Furthermore, increased expression of *NUS1* could possibly promote EMT in breast cancer contributing to invasion and metastases to the brain. It is possible that the silencing of *NUS1P3* in brain lead to silencing of *NUS1* gene that contributes to Mesenchymal to Epithelial Transition (MET). Moreover, *NUS1* expression in CNS and PNS contributes to neuronal proliferation (Huber *et al.*, 2002), and its silencing may provide a selective advantage to supporting cells such as glial cells in metastatic tumour evolution in the brain microenvironment.

CTD-2023M8.1 and *RP11-713P17.4* are long intergenic non-coding RNA (lincRNA) genes. LincRNA are transcribed into RNA without any protein product and are correlated with open chromatic mark such as histone modification sites and epigenetic regulation of transcription, RNA stability, and recruitment of protein complexes (Marchese & Huarte, 2014; Roberts *et al.*, 2014; O'Leary *et al.*, 2015). These are a subclass of long non-coding RNA (lncRNA) with similar biological functions to lncRNA but differing in genomic organizations (Roberts *et al.*, 2014). LincRNA differs from LncRNA mainly because lincRNAs do not overlap with the protein coding genes and are situated between the genes suggesting that the biological roles they played could be attributed to their transcript transcribed independently (Roberts *et al.*, 2014). Similarly, crucial biological functions such cellular growth and differentiation, development and apoptosis are controlled by lincRNAs (Marchese & Huarte, 2014). A type of lncRNA known as HOTAIR (HOX Transcript Antisense RNA) is found to tumour invasion and metastasis due to epigenetic silencing in oral squamous cell carcinoma (Wu *et al.*, 2015). Likewise, LncRNA microarray studies have identified 3 potential lncRNAs associated with invasion and metastases of hepatocellular carcinoma (Gao *et al.*, 2015).

CTD-2023M8.1, a potential candidate metastatic promoter gene is methylated in 26% in metastatic tumours and 63% of unrelated, non-metastatic, primary tumours. It is methylated in 80% in matched primary tumours and none of the BBM samples in individual patients suggesting that the hypomethylation of *CTD-2023M8.1* is a late event during the course of evolution of BBM that takes place only in BBM not in primary breast tumours. In contract, *RP11-713P17.4* is methylated in 11% of the unrelated primary tumours and 76% of the BBM tumours. It is methylated in 30% in the matched primary tumours and 56% of the corresponding BBM samples in individual

patients suggesting that the promoter hypermethylation of *RP11-713P17.4* is an early event during the course of BBM. These two genes, which have RNA transcripts, could possibly regulate genes associated with invasion and metastasis. However, their regulatory functions and the target genes have not been reported before. Functional studies are required to determine the role of these genes in BBM.

MSP in BBM and tumour-free circulating DNA in individual patients has shown some promising results. There is only a little variation in methylation status of BBM compared to serum DNA in individual patients. Moreover, methylation of each patient is identical to BBM and serum DNA in the majority of genes in a panel of six genes *i.e.* *CCDC8*, *MIR3193*, *CTD2023.8*, *FOXD1*, *RP11-713P17* and *MIR124-2* (Table 7.7). Methylation in BBM is the same as in serum DNA in patients 1, 2 and 5 in a panel of genes. In addition, methylation in BBM is >75% identical to serum DNA methylation in patients 7, 8,10 and 12. Similarly, methylation in BBM in patients 6, 11 and 13 is 67%, 50% and 50% identical to serum DNA respectively. These findings suggest that these genes could be used as the basis for a panel of non-invasive prognostic biomarkers to predict the risk of BBM in patients with breast tumours.

These findings merit further investigations to develop these potential biomarkers in clinical settings. The methylation status of this panel of genes could possibly be used during the time of diagnosis of primary tumours either in a serum of patients (non-invasive method) or in primary tumours if surgery has been carried out. In cases where these markers show aberrations in methylation status suggesting the possibility that primary tumour cells have already metastasised, some adjuvant therapies such as demethylating agent treatment or other suitable epigenetic/genetic therapies could be used to prevent or to slow down the growth of the metastases. Similarly, in cases where

there are no aberrations in methylation of these genes at the time of diagnosis patients could be followed up at regular intervals to investigate the possibility of breast tumour metastases.

There is a need to investigate these genes in a large number of clinical specimens. More primary tumours-BBM pairs from individual patients including normal tissues are crucial to investigate these genes. Similarly, it is important to follow up if those unrelated patients with primary breast tumour (a cohort of unrelated primary breast tumours used) developed brain metastases. A number of genes from a panel of genes, which were methylated both in BBM, and unrelated primary tumours from the patients (section 4.2 and section 7.2), could possibly have a role in breast cancer metastases to other secondary sites such as bone, liver and lungs. In this regard, follow up studies in those unrelated patients might provide further information about the genes involved in breast cancer metastases to other sites than the brain.

7.4 Conclusion

Analyses of Illumina HumanMethylation 450K BeadChip array on 24 BBM samples in comparison to the normal breast tissues and primary breast tumours from TCGA identified 7 genes (*MIR3193*, *FOXD1*, *MTND6P4*, *CTD-2023M8.1*, *RP11-713P17.4*, *MIR124*-2 and *NUS1P3*) dysregulated in BBM, which showed statistically significant differences in methylation in BBM compared to primary tumours and normal breast tissues. Three of these genes (*RP11-713P17.4*, *MIR124*-2 and *NUS1P3*) were hypermethylated in BBM compared to primary breast tumours and normal breast tissues whereas other four genes (*MIR3193*, *FOXD1*, *MTND6P4* and *CTD-2023M8.1*) were hypomethylated in BBM compared to primary breast tumours and normal breast tissues.

The genes identified included two microRNAs (*MIR3193* and *MIR124-2*) and four non-coding RNA (ncRNA) genes such as novel long intergenic non-coding RNA (lincRNA) genes (*RP11-713P17.4* and *CTD-2023M8.1*) and two novel non-protein coding genes or the pseudogenes (*MTND6P4* and *NUS1P3*). This suggests that these genes are important regulatory elements and their epigenetic aberration (DNA methylation) contributes to the dysregulation in crucial steps in genetic network during evolution of BBM. Similarly, methylation analyses of these genes in BBM and their corresponding primary tumours in individual patients showed that dysregulation of *RP11-713P17.4*, *MIR3193*, *MTND6P4* is an early event whereas dysregulation of *FOXD1* and *CTD-2023M8.1* is a late event during the process of BBM. In addition, methylation analyses of these genes including *CCDC8*, *BNC1* and *GALNT9* (from section 4 and 5) in tumour free circulating DNA in patients' plasma showed that the methylation status of serum DNA corresponds to BBM in individual patients suggesting that these genes could possibly be potential prognostic biomarkers for BBM or as therapeutic targets.

CHAPTER 8

Final Discussions and Conclusions

8.1 Final Discussion

Cancer metastasis, an extremely complex and multistep process is attributed to 90% of the deaths by cancers (Valastyan & Weinberg, 2011). These multistep cascades consist of various layers of selective pressures to metastasising tumour cells not limited to migration and invasion (Spano *et al.*, 2012). This implies that there should be a complex genetic and epigenetic dysregulation, which possibly contribute to the metastatic cascades favouring metastasising tumour cells during the process of metastatic evolution (Valastyan & Weinberg, 2011). Due to this complex nature of the disease, there is a need of preclinical studies and therapeutic intervention to control the metastatic process (Sleeman, 2012). However, due to the cost and other challenges associated with metastases such as complexity of the diseases, there is an urgent need to identifying and developing biomarkers to predict the early risk of cancer metastases. Moreover, the presence of micrometastases undetected at the time of diagnosis of primary tumours further intensifies an urgent need for developing biomarkers to predict the surviving micrometastases into the new microenvironment. There have been significant advances in understanding tumour biology especially in terms of primary tumours, however, information on primary tumours has not been adequate to make a decision relating to metastasis due to the differences between genetic and epigenetic dysregulation in primary and metastatic tumours (Guo *et al.*, 2011).

The aim of this study was to identify epigenetically dysregulated genes in BBM that could be used as prognostic biomarkers in BBM. Breast cancer is a heterogeneous population of cells as a result of unstable genetic dysregulation and variations arisen as the tumour cell divides (Minn *et al.*, 2005a). Therefore, it is not surprising that primary breast tumours, like other metastasising tumours, consist of subpopulations of cells with higher metastatic potential to disseminate into distant sites *i.e.* bone (Kang *et al.*, 2003), lung (Minn *et al.*, 2005a), brain (Guo *et al.*, 2011), liver (Minn *et al.*, 2005b) and relatively rarely to adrenal glands (Liu *et al.*, 2010). A previous study identified a group of genes (*IL11*, *CTGF*, *CXCR4*, and *MMP-1*), which are overexpressed in a subpopulation of breast cancer cell lines (MDA-MB 231) metastasising to the bone (Kang *et al.*, 2003). Similarly, Minn *et. al* (Minn *et al.*, 2005a) used expression profiling techniques in a breast cancer cell line (MDA-MB-231) and identified 18 genes, which are uniquely dysregulated in breast to lung metastases. In addition, a group of genes were identified that provided a selective advantage to the primary breast tumours as well as to the metastasised tumour cells in the lung microenvironment (Minn *et al.*, 2005a). Furthermore, gene expression studies have identified *COX2*, *PTGS2*, *HBEGF* and *ST6GALNAC5* as involved in the process of BBM (Bos *et al*., 2009). Frequent large chromosomal gains (1q, 5p, 8q, 11q, and 20q) and deletions (8p, 17p, 21p and Xq) were reported by Salhia *et.al* (Salhia *et al.*, 2014) in BBM based on their integrated genetic and epigenetic (27K methylation array) studies. A number of genes overexpressed (*ATAD2*, *BRAF*, *DERL1*, *DNMTRB* and *NEK2A*) and deleted (*ATM*, *CRYAB* and *HSPB2*) were reported a study including the genes (*AURKA*, *AURKB* and *FOXM1*) enriched in cell cycle and G2/M transition pathways (Salhia *et al*., 2014). However, the identified genes were based on only comparison of unrelated primary breast tumours and BBM with non-neoplastic brain and breast tissues. Furthermore, analysis of BBM

and their corresponding primary tumours in individual patients is crucial to identify unique events in BBM compared to primary tumours (Salhia *et al*., 2014). In this regard, due to the limitations of the Illumina HumanMethylation 27K BeadChip array and lack of matched tumours from the individual patients, the study was able to identify only basic epigenetics and genetic events in primary breast tumours and BBM (Salhia *et al*., 2014). Moreover, comparison was made between the normal brain and BBM (unrelated tissues), not BBM and primary breast tumours. Therefore, despite advances in breast cancer diagnosis and therapeutic advancements in recent years, roles of epigenetic determinants such as DNA methylation on breast cancer metastases to distant sites including the brain still remain elusive. The current common predictive methods for BBM include factors that increase the risk of BBM *i.e.* age, ethnicity, receptor ER/PR status, HER2 receptor status and germline mutation in the BRCA1 gene (Salhia *et al.*, 2014).

The genetic dysregulation responsible for the BBM may be common to the genetic dysregulation responsible for one of the primary tumours that readily metastasises to the brain such as lungs, breast, melanoma and RCC (section 2.2.1, section 4.1). Based on these hypotheses, a comprehensive literature review was carried out to screen genes, which were methylated in either lungs, melanoma, RCC but not in primary breast tumours. Investigation of methylation status of a long list of 78 genes in BBM identified two candidate metastatic suppressor genes (*BNC1* and *CCDC8*), which were frequently methylated in BBM and infrequently methylated in unrelated primary tumours (section 4).

Similarly, bioinformatic analyses of 450K methylation data of primary lungs and breast tumours from the TCGA were carried out to identify genes that may be dysregulated in BBM (section 5). Investigation of methylation status of 5 genes in BBM generated from the TCGA data analysis identified *GALNT9* as a potential metastatic suppressor gene frequently methylated in BBM and infrequently methylated in unrelated primary breast tumours.

Study of *in vitro* metastatic potential (wound healing assay and invasion assay) for *BNC1*, *CCDC8* and *GALNT9* in breast cancer cell lines showed that the downregulation of *BNC1*, *CCDC8* and *GALNT9* increases the migratory and invasive potential of MCF7, T47D and MDA-MB 231 respectively.

Genome-wide methylation analysis identified 7 candidate genes of which, three genes (*MIR124-2*, *RP11-713P14.4* and *NUS1P3*) are candidate tumour suppressor genes whereas four genes (*MIR3193*, *MTND6P4*, *FOXD1* and *CTD-2023M8.1*) are metastatic promoter genes. Taken together, this entire project identified 10 candidate genes (Figure 8.1) of which, 6 were metastatic suppressor genes (*BNC1*, *CCDC8*, *GALNT9*, *MIR124-2*, *RP11-713P14.4* and NUS1P3) and 4 were metastatic promoter genes (*MIR3193*, *MTND6P4*, *FOXD1* and *CTD-2023M8.1*). Furthermore, candidate genes identified by genome-wide methylation analyses includes two microRNAs (*MIR3193*, *MIR124-2*), two long intergenic non-coding RNAs (*RP11-713P14.4*, *CTD-2023M8.1*) and two non protein-coding genes (*MTND6P4* and *NUS1P3*).

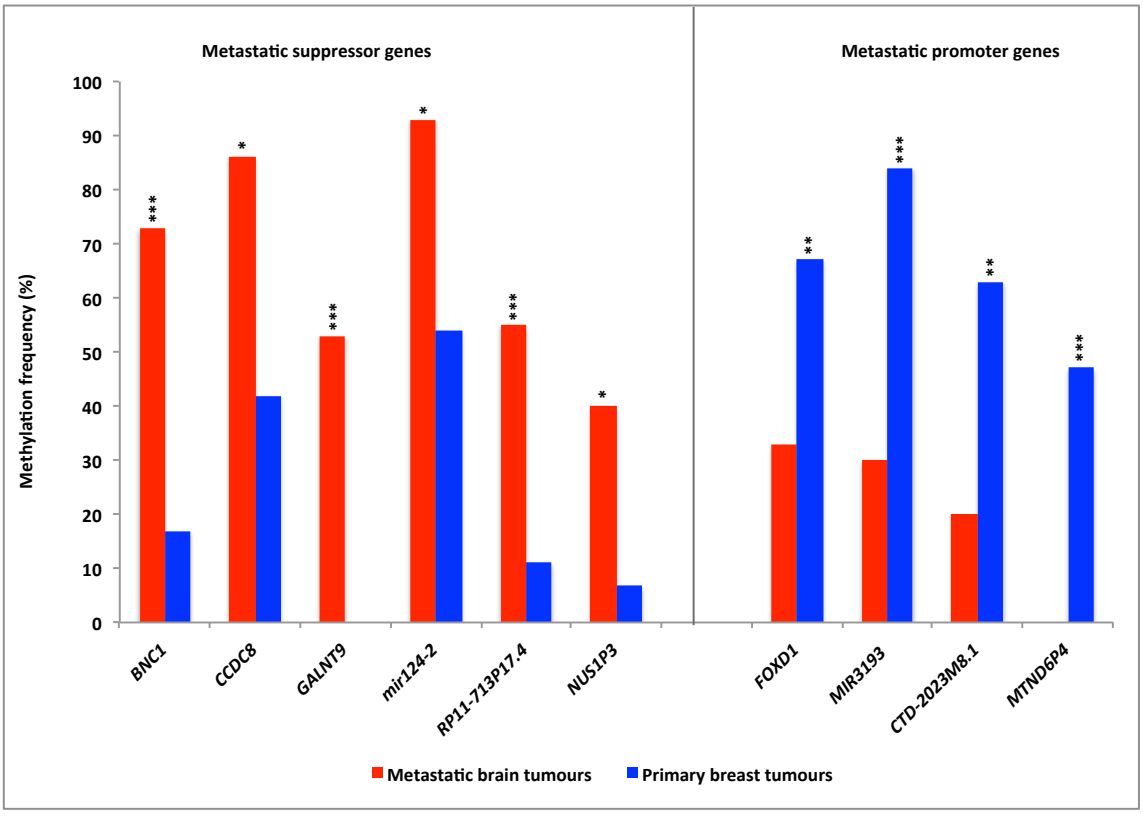


Figure 8.1: A total of 10 genes (candidate prognostic markers) were identified as dysregulated in BBM and are differentially methylated in BBM compared to primary breast tumours. Six of these genes are candidate metastatic suppressor genes hypermethylated and silenced in BBM samples whereas four are candidate metastatic promoter genes hypomethylated and expressed in BBM samples. The difference in the methylation status of these genes in BBM compared to primary breast tumours is statistically significant (*: p≤0.05, **: p≤0.01, *: p≤0.001).

We hypothesised that epigenetic dysregulation of BBM associated genes would either occur as a) early events or as b) late events in tumour evolution (section 2.1). Consistent with our hypothesis, the genes identified in our study are either dysregulated early in the primary breast tumours providing selective advantage to the tumour cells to metastasise to the brain (*CCDC8*, *RP11-713P14.4*, *MIR3193*, and *MTND6P4*) or are dysregulated late in the process of BBM which, contribute to the survival of metastasize tumours in the brain (*BNC1*, *GALNT9*, *FOXD1* and *CTD-2023M8.1*). It is important to note that both early and late methylation events appear similar in the initial analyses using unrelated primary tumours and BBM. However, a comparison of primary breast tumours and BBM from the individual patients have revealed whether specific gene methylation occurs early or late in the process of tumour evolution.

The investigation of methylation status of the genes in tumour free circulating DNA in patients' plasma using MSP was another exciting aspect of this study. Methylation status of a panel of 6 genes (*CCDC8*, *MIR3193*, *FOXD1*, *CTD-2023M8.1*, *RP11-713P17.4* and *MIR124*-2) together in BBM and in patients' plasma was identical in up to 90% of cases (gene dependent) for each patient. This initial screening using patients' plasma suggests that these genes could be used as a panel of prognostic biomarkers for BBM. However, investigation of methylation in many more serum samples is necessary to validate the efficacy of these genes as biomarkers in a clinical setting.

In addition to the ten candidate prognostic biomarkers, this study has identified a number of candidate genes, which deserve further investigation to explore their roles in breast cancer metastases to distant organs. For instance, from the panel of 20 genes frequently methylated in BBM (section 4), 18 genes were frequently methylated both in

primary breast tumours and in BBM. In addition, there were five metastatic suppressor genes from the genome-wide analysis, which were frequently methylated both in primary tumours and BBM. It is possible that these 23 genes are silenced both in primary breast tumours and in BBM suggesting their role in primary tumours as well as in breast cancer metastases to other sites including brain. Moreover, these genes could possibly be passengers for BBM, and a number of them could be specific to other distant sites. These genes were excluded from the further analyses due to the limitation of the study at this point.

Nguyen *et. al.* (Nguyen *et al.*, 2009) proposed the possibility of three different classes of genes involved in the process of metastases *i.e.* 1) metastasis initiator genes, 2) metastasis progression genes and 3) metastasis virulence genes (section 4.3.). In this regard, there is a possibility that these 23 genes, may fall in one of these three categories. It is important to note that these 23 genes primarily play a role in primary tumours and may be only a passenger for the metastatic process. However, frequent methylation status of these genes both in primary tumours and BBM merits further investigation on methylation of these genes in breast to lung, liver and bone metastases. Moreover, follow up of these patients for their breast cancer recurrence in any secondary sites and investigation of methylation of these genes on those sites could possibly confirm if these genes contribute to breast cancer metastases to any other sites of the body.

Likewise, it would be worth investigating the methylation status of these 10 genes, (dysregulated in BBM) in breast to lung, bone or liver metastases. It is possible that the silencing of these genes may contribute to survival of the metastasised tumours in lungs, liver or bone as micrometastases. Furthermore, it may also be possible that methylation

of some of the genes may contribute to brain metastases in a later stage from the other secondary sites to which breast metastases initially had taken place. Therefore, a number of candidate genes dysregulated in primary and BBM also suggests that these genes could play an important role in one of the steps of the metastatic cascades of breast cancer metastases to BBM.

8.2 Final Conclusions

This study has identified ten candidate genes differentially methylated in BBM compared to primary breast tumours. These candidates were identified using three independent approaches *i.e.* broad literature review, bioinformatic analysis of TCGA data and genome wide 450K methylation analyses. A broad literature review identified *BNC1* and *CCDC8*, analyses of the TCGA data identified *GALNT9* and the genome-wide methylation analyses identified *MIR124-2*, *RP11-713P17.4, NUS1P3*, *MIR3193*, *MTND6P4*, *FOXD1* and *CTD-2023M8.1*. Methylation analyses of these genes in unrelated primary breast tumours and BBM samples revealed that six of these genes (*BNC1, CCDC8*, *GALNT9*, *MIR124-2*, *RP11-713P17.4* and *NUS1P3*) are metastatic suppressor genes hypermethylated and silenced in BBM where as the other four (*MIR3193*, *MTND6P4*, *FOXD1* and *CTD-2023M8.1*) are metastatic promoter genes hypomethylated and expressed in BBM samples. Similarly, methylation analyses of these genes in BBM and their corresponding primary breast tumours from the individual patients revealed that the dysregulation of *CCDC8*, *RP11-713P17.4* and *NUS1P3* occurs early in primary tumours whereas the dysregulation of *MIR3193*, *MTND6P4*, *FOXD1* and *CTD-2023M8.1* occurs late only after the tumour cells has left the primary tumours. Study in the *in vitro* metastatic potential of *BNC1*, *CCDC8* and *GALNT9* in

breast cancer cell lines showed that the downregulation of *BNC1*, *CCDC8* and *GALNT9* increases the migratory and invasive potential of respective breast cancer cell lines used. In addition, methylation status of a panel of six genes (*CCDC8*, *MIR3193*, *FOXD1*, *CTD-2023M8.1*, *RP11-713P17.4* and *MIR124*-2) in patients' serum has showed that these genes could be used as non-invasive prognostic markers to predict the risk of BBM. However, analysis of many more patients' serum samples is necessary to validate these markers before taking these candidates into a clinical setting.

Appendix

Appendix A

Appendix A1:

Breast to Brain metastases			Lung to brain metastases			Melanoma-Brain metastases	RCC-Brain metastases	Colorectal Carcinoma-Brain metastases
BM1	BM12	BM22	LM1	LM11	LM21	MM1	RM1	CM1
BM2	BM13	BM23	LM2	LM12	LM22	MM2	RM2	CM2
BM3	BM14	BM24	LM3	LM13	LM23	MM3	RM3	CM3
BM4	BM15	BM25	LM4	LM14	LM24	MM4	RM4	CM4
BM5	BM16	BM26	LM5	LM15	LM25	MM5	RM5	CM5
BM6	BM17	BM27	LM6	LM16	LM26	MM6	RM6	CM6
BM7	BM18	BM28	LM7	LM17	LM27	MM7	RM7	CM7
BM8	BM19	BM29	LM8	LM18	LM28	MM8	RM8	
BM9	BM20	BM30	LM9	LM19	LM29	MM9		
BM10	BM21	BM31	LM10	LM20				
BM11								

Appendix A1: A list of brain metastases from breast (BM1-BM31), lungs (LM1-LM29), melanoma (MM1-MM9), Renal Cell carcinoma or RCC (RM1-RM8) and colorectal (CM1-CM7). Genomic DNA was extracted from these tumours. BBM samples were used for the methylation analyses.

Appendix A2

Primary breast tumour code	Primary breast tumour code	Primary breast tumour code	Primary breast tumour code	Primary breast tumour code
BP101	BP120	BP139	BP158	BP178
BP102	BP121	BP140	BP159	BP179
BP103	BP122	BP141	BP160	BP180
BP104	BP123	BP142	BP161	BP181
BP105	BP124	BP143	BP162	BP182
BP106	BP125	BP144	BP163	BP183
BP107	BP126	BP145	BP164	BP184
BP108	BP127	BP146	BP165	BP185
BP109	BP128	BP147	BP167	BP186
BP110	BP129	BP148	BP168	BP187
BP111	BP130	BP149	BP169	BP188
BP112	BP131	BP150	BP170	BP189
BP113	BP132	BP151	BP171	BP190
BP114	BP133	BP152	BP172	BP191
BP115	BP134	BP153	BP173	BP192
BP116	BP135	BP154	BP174	BP193
BP117	BP136	BP155	BP175	BP194
BP118	BP137	BP156	BP176	BP195
BP119	BP138	BP157	BP177	BP196

Appendix A2: List of unrelated primary breast tumours used in methylation analyses.

Appendix A3.1:

Gene	Primer	Sequence (5'-3')
ABCB1	F	GTG AAG TTT TTT GGT AAG TTT ATG GGG ATT
	IF	GTA GYG TTT AAA TYG TAG TGG TAT TGG ATT
	R	AAC CAA ATA CAT GAA CCT CAG GCR CRC TAA
AK5	F	GAG YGG GYG AGT GYG TTT TTA TAA GTT
	IF	AYG GGG YGY GGT TGT TTT AAT TYG AAT T
	R	AAA ACC ACA ACC CCC RAA AAA AAA CRA AAA
ALDH1A3	F	TAG TTT TTT GTY GGG TYG GGT GTT TTA GTT
	IF	GTT TTT YGG GGA GTT YGT TTA TAG GTA GTT
	R	TTT CCA CRA CCC CRT TAA CRA TAA CCA TAA
ANK3	F	TG GGG ATG GTT AAT GAA ATT TTA TTT AGT AGT T
	IF	G AGT TGT TTT TGT TGT TTT TAG TTA TTT TAA GGT

	R	CCT CRA ACC TCC AAA CRT AAA ACT TCT TAA
ATM	F	GGA TAG ATT GGG TYG TAT AYG ATT GAA TTT TT
	IF	GAA GTY GTT GYG TTG TTT TYG YGT TTT TAT T
	R	ATA AAA ACC CCA AAA CCC RAA CCR AAA A
BNC1	F	GGT YGA ATT ATT TTT TGA GAA GAG YGT TAG AGA AT
	IF	GAG AGG TAA ATA TYG ATA YGT TTG TGT TTT TTA TT
	R	CCA AAC RCC CAA ACT ACR CAA CAC RAT AA
BOLL	F	GGY GAT TTT AAG GTT GTT GAA GGT TGA AAT
	IF	TGG TTT YGT TTY GGA TTT TTT TAG AGG TTG TGT
	R	CTC TCC TCC ACT TAA AAA AAA TCR ACC TCT AA
BVES	F	GGT AGT GTT TTG TTT AYG ATG GAG GGT T
	IF	GTA TYG TTT GGG ATT GTT AGT TTY GGT TAA GTA TT
	R	TCR CCT CRC TAC AAA ATT CTA TCC TTA TTT AA
CCDC8	F	GTT ATT GTY GTT TYG TTG GGA AGT AAT GGG TT
	IF	GGT YGT TTY GAA GGG AGT GGA TAG TT
	R	CCC AAA ACT CAA AAC CCA CTC CAA ACT AA
CDKN1A	F	ATA GTG GYG TAA AGG ATT TGA ATT TGG GAT T
	IF	GGG YGG YGT AGA TTG GAT TTT AGA GTT ATT
	R	CTC RAC TTA CRC ACA CRA TAT CTC TAA ATA
CD44	F	TTA GYG GGA AAG GAG AGG TTA AAG GTT GAA TT
	IF	TGG TGT AAG GTT TTA YGG TTY GGT TAT TT
	R	CCR AAC CTA ACA AAA ACT AAA ATC CRC TAA
CLDN1	F	GYG TGA AAY GTT TTA TAG GAG YGA GAA GAT TT
	IF	GTA GGG AYG TAG TTT TGG TGT TTG GTT T
	R	CC TCC ACT AAA ACA AAA CAA TAC TAA CRA TAA
CLDN3	F	GTT TAG TTA GGT TTA AGG GTA TTT TTT GGG TAT T
	IF	GGT TTT TGG GTT TAG TAA TTT TTG GGT TGG AT
	R	AAC CTA AAA ACA ATA ACR TAA CCC CTA CC
CLDN4	F	GTY GGT TGG AAG GAA TTG GTT TGT TTA T
	IF	GGT TTT ATT TTT TGA TTT AYG GTG TAA AGG TGT AT
	R	C CAT AAA CCC TCC CAA ATA ATC TAC RAA ATA
CLDN5	F	GAY GGG GTG GGG GGT TTT TGT AAA TT
	IF	GGA GTT GGT TYG AGT GGA AGA GAT T
	R	CC AAC TCC TAC CRA AAA ATA CCC TCT TTA
CLDN6	F	GTG GTG AAG YGG AGT TTT AAG TTT TGT TTT ATT
	IF	GGG TGA GTT TGG AYG TTT GTT AGT T
	R	CCR AAA AAC CCT ATC ACC TCR AAA ACT TAA
CLDN7	F	GGA AAT TTA GTT YGG AGG GGT TGA T
	IF	GAG TTT GGG ATT TTT GGG GAG TTA T
	R	CCC AAA ATA TCC TAA ACT ATA AAT CCR AAA CTA
CLDN9	F	GTA GGT YGT AYG TGT TTT TTG TGT TAT TGT T
	IF	GTT GGT TTT GTT TGG TTT TTT GGT GGT T
	R	CAA AAA AAC CCC CCR CCC AAC ATA AAC AAT A
CLDN11	F	GG TYG GAG YGG AAT GGA TAT TTA GAG AT
	IF	GAT GAG AGA GGG GTT ATA AGA AGA GAA ATT

	R	CTC ACR ACT ATA CCT AAA CCT AAA CTA AAC AAT A
CLDN18	F	GGA GAG TAG ATA AGT GTT TTT AGG GTT
	IF	GAA TTA GGA GYG AAA TTG AGT TAT TTA AGG AAA AT
	R	CAA CTC CTC CAA AAC CCT TCR TAC TAA
CLDN23	F	GA TAY GTT TTT GGA TTG GGT GGG TTT TTG TAT
	IF	GYG AGT GGT TGT AAG TTG TAG AGT AAA TT
	R	CAA CCR AAC RAC TAA ACC ACR CCT AA
CMTM8	F	GYG GAG TTT TTG GAG GTT TGG GTT TTT
	IF	GTT TTY GAT TYG TTT AGG TGG AAG GAA AT
	R	CTC CCR AAC TAC RCT AAA ACC CTA AAA A
COL14A1	F	GTATATAGTATTTGAGAATAGGAGGGTTYGAGAT
	IF	GGGTTTGGTTTGGGAGTTTGTAGTAGT
	R	CTAAAACTACCTCCRCCTCTCCAACTAA
CSNK1A1	F	GGT GAT TTT TTT AAA TAT TAA GGT YGG GGG GTT
	IF	GGG TTY GAG GYG AGT AAT AGG TAA TAT
	R	CTA CCA CCA CTA CCR CCR ACT CC
DAPK1	F	GAA GTT TTY GTT TAG AYG GTA YGY GTT TTA T
	IF	TTG YGG TAA YGG TGG TAT TTT YGY GTT ATT
	R	CAA CTT TTA CTT TCC CAA CCA AAA CRC RAT AA
DGKI	F	GTT TTA YGG TGT AGG ATT TAG GGY GYG GTT
	IF	G TTT TGG YGT TTT TAG TTY GGG GTY GTA GTT
	R	CAA ACR ACR CCR CTA CAA AAA CCR AAC RAA CTA
DLC1	F	TTG YGT GTT TGT TTG YGG YGT TGY GTT TTT
	IF	GGT TAT TTG GAT TGT TTT TGA AGG GAA GAT T
	R	ACC RCA AAC TAT CRC CTA CCT TCA AAA AA
DSP	F	GTY GGG TTA YGT ATT TTT AGT TYG AGA GGT T
	IR	GAG GTT TTA GYG TAG AGY GTA GTA GTT
	R	CTA ACA CTA AAC RCC TAA ACC ACA CCT AA
FBLX14	F	GAT AGT TGT YGT TTT TTT GTG GYG TTT AT
	IF	GGG TTA AAA ATT AGG GGT TAY GGG ATT
	R	CCR CCR CCC AAA CRA AAA AAC CAA A
FBN2	F	TGY GTY GGT TTT TTA GTT YGG GTT GAG TT
	IF	TGT GGA YGG TYG YGT TTT TAT TGA TAG T
	R	ATT CCC RTA CAC TCC RAA AAC RAA TAT TAA
GALNT9	F	GGT GAG GTG AGG TYG YGT TAT TTT GTA T
	IF	GGT AGG ATA GGA GTY GGT ATA GAT AT
	R	CAC CRT CAA CAA AAT TCR AAT CTT CCT AA
GATA5	F	TGY GGA GTT TTA GYG GTT YGG GTT TTT
	IF	GTY GTT TTT TTT GTT YGT TGG TGA GTG TTT T
	R	CAT CCT CCC AAA CRT AAT CTT AAC CTA CAA A
GREM1	F	TAG TAG AGG TTG GTT ATT TGT YGT TYG TTT
	IF	GYG TAG GTA AAT AAA YGT TAA GTT GGG GTA TT
	R	TAA CRC CRA ACA ACR ACC ACC RAC TAA
HK2	F	GTG GGA GTT TTT TAT ATG ATT TYG AGA TGT T
	IF	TAT AGG GYG TGT GTT GGT TTA GAG GTT
	R	CTC CTA CRC CRA AAT TTC ATA CAA CAA TAT AA

HOXB13	F	GTT GGG GTA AAG TAT TTT YGT AGT TTT TGT T
	IR	CTC CCA ACA AAC CTT CRA TAT CCT TAA
	R	CAA ACA TCA ACR TAA ACR CCR CTA AAT AAC TAA
HOXD3	F	GTG GYG GTT TYG GGT GAT TTT T TA ATT TAG T
	IR	CAA AAT CCC CCT CTT ACA TCT ACC CTA TA
	R	CAA CRA ATA ACC ACC CCC CRT CAA ATA
HOXD10	F	GAG GTT TTT TYG TAT TAG AGG TTG GTT T
	IF	GAT AAG YGT AAT AAA TTT ATT TYG GTY GAG GTT T
	R	CTC TCC ACC TAA AAA ACT TTA ATA AAC TCC TAA
ICAM5	F	GYG GGG TTY GGA GTT TAG TTT ATT ATT ATG GT
	IF	GTA GGG GTT TGG GAA GGG TAA TTT AGT ATT T
	R	CCC CAA ACC CAA AAC AAC CCA AAA A
IGFBP3	F	YGA TAT YGG TTY GTY GTA GGG AGA TTT TAT T
	IF	GGG TGT TGA GTT GGT TAG GAG TGA T
	R	CAA CRC CCA ACC RCA ATA CTC RCA TCT AAA
KLHL35	F	GYGGAGYGGTATTTTTATYGAYGTGGTGT
	IF	GTTTAGYGYGGGTAGYGTTTATTTTYGTAGT
	R	CRTAACRCRCCACCTCRACRAAACCTA
KRT7	F	GGG ATG GYG TTT TTG TTT ATT TYG GAT T
		GGT TTG GTA GTA GAG AAA GGT GGT T
		CAA CCR CAA CRA AAC CAA CAA ACT CTA A
KRT18	F	GTT TTG AAA GTA GTT TYG AGG GTT AAT AAT AT
	IF	GTT GTY GTG TTT ATG TTY GGT TGG TTA T
	R	CCT AAA CAA AAC CCA AAA ACC RAT AAT TAA TA
KRT19	F	GTT TTT GGT TTT TGG GAG GGG AGG GAA TTG ATT
	IR	GGG GAG AGA GTT TAT ATT TGT TTT TAG GAG TT
	R	CAC CTT ATC CAA ATA AAA AAC CAA ACR ATC RTT A
KRT28	F	GAT TGA GTG TGA ATT TGT AAT GTT TTT AAA GTT AT
	IF	G AAA ATC TAC AAA ATA CRC CTC ATC TCR ATA
	R	CT AAC CTT CRC TAT CAC CTA CCR TAA
KRT72	F	GG ATT TAT GGG GAG TTT AYG GTG GAA T
	IF	GGT TTT AAG GGA TTT AAA AGG TGT TGT YGA TAT AT
	R	CTA CCC CCA AAA CAA AAA AAA CTC TTA CTA
KRT81	F	GTA GGY GGA AGA AAA AGA TGA ATT TTY GGG AT
	IF	GAT TAT TAG GAA AGG TTA TAG AGA GAG ATT
	R	CTC CCR CTA AAC RAA TCT AAA AAC CTC TAA
KRT83	F	GGA AAT TTT AGT TGT GTT TTT GTT TGY GGG TT
	IF	CT ATT AAA AAA CTT AAT CTA CTC CTT CTC CTC CTA
	R	CCA AAA CRA CAC CCA CCT TAT CRA TA
KRT85	F	GGA TGY GGG GTT ATT AGG AAT TTT AGT
	IF	GTG GTT TTT AAA ATT GGT AAT YGT TGT TGT ATT
	R	CTT AAT CTA CTC CTT CTC CTC CTA
KRT86	F	GGT TTA TAG GGT GTA AGT AGT GAA YGT T
	IF	CCA TCC TTA ATA CCC RAC CCT AAC TAT A
	R	CTA TAT ATC TAT ATA CCT CRC CCA TCA AAC TAA

MIR23B	F	GGA GGA TGA GAG ATT ATT TGA TTT TAA GGA TGT AT
	IF	GTT GGG AAT GTT AGT TAT GTA TTG AGT TT
	R	CTC RAC TCC TAT TCC TAC TAA ACT AAA CCA ATA
MIR34a	F	GAG ATA ATA GGT TTT GAT TYG GGA TAG ATT TTA T
	IF	GGG TTT TTT TTT TTT TTT AGG YGA GAG GTT
	R	CCT AAC CTC RAC CTA AAA CRC AAA TAC RAT A
MIR34b	F	GAG TTA GTT TTA GGG TTT GGG TTT GGG AAT
	IF	GTT TGG YGT GAA GGA AGT GGG AGT T
	R	CCA CCA CAA TAC AAT CAA CTA ATA ACA CTA
MIR124-1	F	GA TAG GTA GTY GGA GGG AGC TTT AGA TT
	IF	GG YGT TTG GTT YGG TYG GTY GTT TT
	R	CCC CTC CRC TTT ACA AAA AAA AAC CTA A
MIR127	F	GGY GGG YGG AAG TAG TTT TTT AGG T
	IF	G AGT GGA AGG TGT AGY GGT TTT TGA TTT
	R	C CTA AAT CRA CTC ACC TAC TCC AAA AAA ATA
MIR433	F	GGA GTA GAA AGY GTA GTA GGT TTT TTT GT
	IF	GTT TTT AAG TAG AAT GGG TTT TGG GGT
	R	CCA TCA TAA CAA AAT ACC CTA CCC CTA A
MMP2	F	GTT TAG TYG GTT ATA TTT GGY YGT TGT T
	IF	GGA GGT TGY GTA TTT GGG GTT TTA AAT
	R	CAA ACA ACC CAA AAA ACA AAA CRC CCT CAA A
MST1R	F	AYG TTA GGT GAA GGT ATA GGA GTT AGG TT
	IR	GAG GGT TTG GGT TAG GTT AAG TTT TTT
	R	CCT ATA CCA AAC CTC CRA CRA AAA AAC TAA
NRCAM	F	TAG TTT TTY GAG GYG TYG GYG GTT GGA GTT
	IF	TAG TAG TAG GGT AAG AGG GGG TTG GTT
	R	CAA CTC TAA CRC RAC TAC CCA AAA CCC TAA
OCLN	F	TAT TTT AAG GTT TTA TTY GAA GTA GGY GGA GTA T
	IF	GAG GAY GTG TTT TTT TTA TTA AAG TGT TGA GTG TT
	R	CCR ACT CCT CRA CRC TCT AAA CCT AA
PBRM1	F	GATYGTYGTAGGGATTTTTAATTTTYGGTTTTATT
	IF	AGGGTTAAGGTTTAYGYGTAGGTTTTTATT
	R	CRTAACCRCCACRACTACTACTATTA
PCDH8	F	GTG TTG ATA AGA AAA GGG GGA AAA AGT
	IF	CAA AAA CTA CCC CAA CRC CTC ACA AA
	R	G TTT TTT TAT ATT AAA GGG ATT GTT AGA GGT AGG
PENK	F	AYG GTG TAA TTT TAG TGA TTT TAT GYG GAG AAA T
	IF	GTT TGT TAT TTG GTY GTG TGG GGA GTT
	R	AAA CRA ACC CRA ATA TAC TAA CRA ACR TAA AAC C
PNN	F	ATT TTG YGT TAG TAG YGG GAT TTA GGT ATT
	IF	TAG TTA ATT GGG GAG GGA GAT GAT TGT T
	R	TCT CAT CCA CRT TCT TAA AAC TCT CTT TAA
PTEN	F	GTA GAG TAA GGA GTG AGT TTT AGG TTT TAG TTT
	IF	GYG AGG AGT GGT ATT AGT TTG GGG AT
	R	CTT TCC AAA TTC CCA CTC CCC AAT A
PYCARD	F	TGG TGT AAG TTT AGA GAT AAG TAG GYG TTA TT

	IF	GYG GTT TTT TTT YGG GYT TTA TTT YGT AGG TT
	R	CTT AAA CTT CTT AAA CTC CTC RAC RAT CAA A
QPCT	F	GTT GTG GGT ATA TTT TTA GGY GTA TTT YGT ATT
	IF	ACR ACT CTA ATC CRA TAC RAA AAC CCC TAA
	R	CCA AAA AAA TCA ATC CRA ATA CCT CAA CCT AA
RBP1	F	GTG AGA AGT TAA TGG YGT TTG AAG GAA ATT T
	IR	CCA ATT AAC CAC AAA CRA ACR AAA CRA CTA A
	R	CAA CAT CTT CCA ATA CCC AAT AAA ATC RAC TAA
SDHD	F	AAT GGG ATG TAG TYG GGA TYG AGT ATT
	IF	GTG TAT AGA TTT TYG AGY TGT TTT AGG ATT ATT
	R	CCR CCA TCT CRT TCC TAA AAA CTC AAA A
SFRP2	F	TAT TGT GGG GGY GTA GYG GTT AGG TTT TTT
	IF	ATT TAG GGG AGG GGG TGT AGT TAG AAT TTT
	R	CTA AAT ACR ACT CRA AAC CCC RAA AAA CTA A
STAT3	F	TAG ATA TTY GGT GGT GGT TGT AGG GTT
	IF	GTG ATA ATG TAG GGA AAG GYG TTT TAA TT
	R	ACC AAT CCC TAC TAT CRC TAA AAC CCT TAA
SULF2	F	AGG GAT AGY GGA AAT TAY GGT AGA TAG TAT
	IF	TAA TTT TTY GGG GAG TTT TYG GGY GYG TAT
	R	CRC TTT CTT CTT CTC CTC TCT CRA AAC TAA
TFAP2A	F	GGG TTG GTA GAG TTA GAT TYG TTA AGG TT
	IF	GGG AYG GYG ATT TGT TTT TAT TGT GTT T
	R	CAC CAA CAT CTC ACC TTT TCA TAA AAA ACT AA
TJP1	F	GGT ATR GGT TTT GTR GGT TTT TTT AGT RGT AT
	IF	GTG GAT AAG TTT TTT AAG GAA AGT TTT GGT GTA T
	R	CTA CCR ACC CRA CCC ACT AAA CAT A
TMEFF2	F	GGT TAA AGA GTG TGT TTA ATT GTT TGA AGA ATG T
	IF	GTT TGA AGA ATG TAG TAG AYG GAA GGY GGG TTT
	R	TCC AAC CCR TAA CCT ATT AAT ATA TCC ACC TAA
TNFRSF10C	F	GGG TAG TTA GGG GGA TYG TTT TTT TGT TTT
	IF	GTY GGG TTT TTT TTG TTT AGT TYG GGG ATT
	R	TAA AAA CAA AAC RCC CCR ACC ACC TAA TAA
TNFRSF10D	F	TYG TTG TYG GAA AGA GTT AGT TTT TGT TYG TTT
	IF	GTT TTT TTG GAG GTG TTG GGG GAG ATT
	R	TCT AAT TCC CRA CRC TAT CCT AAC TCC TAA
TSC1	F	GGT TTT TTA TTT AYG YGG GTT TTA GTA GTT TT
	IF	TYG TTG TTT TGA GGT YGT TTA TAG AGA GAT T
	R	CCR AAA AAA AAA AAC TCT TCC ACT CAT AAC TAA
TSPAN4	F	GGT GAT AAG AYG TTT GYG ATT TGY TAG GTT
	IF	GGG GAT AAG GTT TTT TYG AAA GTA GAT TAT
	R	CAT AAT CAT CCC TAT ACR CTA CTT CAC TTA TA
TSPAN13	F	GTT AGY GTG GAT TTT TTT YGA GTT TYG TTT TTT
	IF	GGT YGT YGT TTT GGT TTG GGG TTT T
	R	CCC TAA CAA CTA CTT TAT CCT TAC CTT TAA
UCHL1	F	TTG GAA TAG GAG TTT AGG GAG TAG GTA TT
	IR	AAT AAC RCT TCR TAA ACR TAC TAA AAC TAA
	R	CCT CTC CRC AAA TAC TAT CCC RAC TAA

WIF1	F	AAA GGG AGA YGA AGG GTA ATT TGY GTT
	IF	AGY GTY GGA GGA GGA ATA GGA GTT ATT ATT
	R	CTT AAA CRA CCR CCA CTT AAA AAC RCT ATA A
Primers designed for gene in genome wide methylation analyses		
ZNF808	F	GTT TTA GAT AAT TTT AGG GTT TTG GAA TTA TGT T
	IR	CAC RTC CCA AAA ACA AAA ACR ACR AAA ACC TAA
	R	CCT CTC ACT AAA CTA TCT CAA TTT ACT CTA A
HSPB9	F	GTT TTA GGT TTG GTT TTT AGA AGG TAA GAT AGT T
	IF	GGT TTG GTT TGG TYG TGG ATT TAG TTT T
	R	CTC TAC ATC CRA ATC AAC AAC TAA CTA CTA
DDX52	F	GAG ATY GAG ATG GGA GGA TTA TTT GAG TT
	IF	GYG TTT TAG TTT GGG YGA TAG AGA GAG ATT
	R	CCA ACA ACR AAT TCA TTC CCR AAC TCC TA
EDARADD	F	GYG TGT GTT ATT ATG TAG YGG TAG AGT T
	IF	GTA TTT TAG TTT AGG TTT GTG TAA GGG GAG TTT
	R	CTT CCA CRA ACT CAC TTC CTA CTT TAA
SOX5	F	GGA AGA AGT TGA TAA AGG GAT TTG ATT AAA T
	IF	GAT GYG AGT TTT ATY GAG ATT TAG AGA TAT TT
	R	CAT AAA ACA ACT TCA AAA ACA AAC AAC ACA ATA
RBM23	F	GTG GTY GGG AAG GTT GTT TAG GTA GT
	IF	GTT TAG AAY GGG GAT TTY GAG TTT AGT ATT T
	R	CAA AAC CTC CTC TTC CCR CAA AAT AA
RNF8	F	GGA AAG TTT AAA GTT GTT TAA AAG TGG GT
	IF	GGT TTT AGT TTG GGG YGG TTA GGA T
	R	CCC TAA CCT AAA TTC TAA CTA CCT CTA TA
LMX1B	F	GGA TTG ATA AGT AGG TGA TAG AGG AGT
	IF	CAT ATC CTA CAA ACC CAT TTC CTT TAT CCR TTA A
	R	CTT ACR ACC CTC TAA CRA TCA CTC CAA A
mir124-2	F	GTT TTG GTA GAT GTG TTA GAG ATG AGT
	IF	GAG ATA GGA GTT GGG TTT ATG AGT TAT GAT
	R	CAC AAA CRA AAA CTA CTA CTC TCC AAA CAA CTA
RP11-713P17.4	F	GTY GAG AAG AGT YGA GAA GAG YGG AT
	IF	GTG TYG TTT TTA AGG GGT AGT TTT GGT GTT AT
	R	CCT CAA TCC CRA ACA AAT CAA TTA CCR CTA
NUS1P3	F	GAT TAA GAG GTT AGG AGA TYG AGA TTA TTT TGG T
	IF	GGT GGY GGG AGT TTG TAG TTT TAG TTA T
	R	CCC ATA CRC ACA AAC AAC TTC TCC AAA
MIR3193	F	GAG AGG GYG TGG GYG TTT GTA TTA T
	IF	GTA TTT GGT TAT GYG GGG GTG TTT AAG GAG TT
	R	CAA TAC CTC AAC CTC AAC CCC AAA CTA A
CTD-2023M8.1	F	GGA GGA AAA TAT TGG GTA ATT GTG GGT TAT
	IF	CTT ACC CTC CTA ACC AAA AAA TTC TTA AAC TCA AA
	R	CAT ATA CAA AAA TCT TCA AAT AAC CTC CCC ATT TA
MTND6P4	F	GAA GGA TAT TTG GTT TTA TGG GAG GAT ATA GTT

	IF	GTT GTT ATA GTT GTA AGT AGG AGG ATG ATG TT
	R	CCC CAA TAC RCA AAA TTA ACC CCT TA
FOXD1	F	GGG GAT TTT GTA TTA AGG GAT TGT TTT GTT T
	IF	CTT CCT CCC TAC CCC AAA TCR AAA ATT AA
	R	CTA CCA AAC RTC AAA AAA ACC TCT AAT ACC TAA
C17ORF107	F	GAA AAT ATY GGG GTG GGT TTT AGG AGT T
	IF	CCA AAC TCC TCA AAT TCC TTA TTT TCR TCA TAA
	R	CCA ACC CAT ACC AAA AAA AAC CRT ATT TA
AL2D1	F	GAT TTG TGT TTT TYG GAG AGA TAG AAG TGT TAT
	IF	CAA ACT CAC AAA CRA CTC CAC AAA CTT AA
	R	CAC TCA AAC TCA AAC CCC RAC TAC ATA A
HMGN2P19		GTA TAA TGT GAA GAA AAA GAG GYG AGA AYG ATT
		CCA ACA ACA ACT TTT CCC TTT TTC CCT TTA A
		CTC CAA CAC CTT CAA CTT TCT ATA CCT AA
RP11-734118.1		GAG GGG AGT TTT TAG ATA AAG GYG AGA TAT
		GAA AGG TGA TAG YGT GTT GGT AGT TT
	R	CTA CCC TAC RAA AAA ACA ACT AAA ACC CRA TA
AL662890.1	F	GYG GGT AGT GTT AAA GAT AGG GAT GTT
	IF	GAT TTY GGG GTT AGA AGT TGA GGA T
	R	CCC CTA CAC CTA AAA CTA TTT ATT TCC TTA

Appendix A3.1: A list of CoBRA primers designed to amplify promoter region of candidate genes (chapter 4, 5 and 7). Y: C or T in a forward sequence; R: G or A in a reverse sequence.

Appendix A3.2

Gene	Primer type	Primer	Primer sequence
BNC1	USP	F	G GTG GTT TTT TGG GTG GTG AAG TAG T
		R	CAA AAT AAA CAA CTC CCC AAA CAC CC
	MSP	F	CG GTT TTT CGG GCG GCG AAG TAG T
		R	CGA AAT AAA CGA CTC CCC AAA CGC C
CCDC8	USP	F	GGT GAG GAG GAA TTT TTT GTT TGT GTG T
		R	C CAA AAC ACT AAA AAA CCC CAA CCC AAA
	MSP	F	C GAG GAG GAA TTT TTT GTT CGC GCG T
		R	CGA AAC GCT AAA AAA CCC CAA CCC GAA
CTD-2023M8.1	USP		GTG ATA AGT ATG GTT GTT AGA TTT AAG ATA TGG
			CTC ACT ATA CCA TAA TCT CAA CTA AAT CAA TA
	MSP		GCG ATA AGT ATG GTT GTT AGA TTT AAG ATA CGG

			CTC GCA ATA CCA TAA TCT CGA CTA AAT CAA TA
FOXD1	USP	F	GTT AGT GTG TTT GAG GAA GAA GGT AGG AAT TT
		R	CCC ACA AAA CCA AAA AAC ACA CCC AAA CTA
	MSP	F	GC GCG TTC GAG GAA GAA GGT AGG AAT TT
		R	C GCA AAA CCA AAA AAC GCG CCC GAA CTA
GALNT9	USP	F	GTT TTG GGT AGT TTT TTG TGT TTT GGG GGA
		R	CAC AAC TAC AAC AAA AAC CTC ACC CAC AAA
	MSP	F	CG GGT AGT TTT TTG CGT TTT GGG GGA
		R	C GAC TAC AAC GAA AAC CTC ACC CAC GAA
MIR124-2	USP	F	GTA TGT GGT GAA TGT TAA GAG TGG AGT T
		R	CCT CAA AAA TTT ACC TAC AAA TTT CCA CAA CTA
	MSP	F	CGC GGT GAA TGT TAA GAG CGG AGT T
		R	CGA AAA TTT ACC TAC GAA TTT CCG CGA CTA
MIR3193	USP	F	**G TGT TTT TTT GTT GTG TTT TTT GGG GAA G**
		R	**CTT CAA CAA CAC AAA ATA CAA TCA TCC ACA AA**
	MSP	F	**CGT TTT TTC GTC GTG TTT TTT GGG GAA G**
		R	**CGA CAC AAA ATA CGA TCG TCC GCA AA**
MTND6P4	USP	F	TGT TAT TGG TGT GAA GGT AGT GGA TGA TTT
		R	CT CCA CAT AAT AAA ACT TCA ACT CAC TTC TTA A
	MSP	F	CGT TAT TGG CGT GAA GGT AGC GGA TGA TTT
		R	CCG CAT AAT AAA ACT TCG ACT CAC TTC TTA A
RP11-713	USP	F	GAA GTT TTT TGG TTG AGT TGT GGT ATT TAT GT
		R	CTA ACC TAA AAA TAC CAA AAC AAG CCT AA
	MSP	F	GTT TTT CGG TCG AGT TGT GGT ATT TAC GT
		R	CC TAA AAA TAC CGA AAC GAC CCT AA

Appendix A3.2: Methylation Specific PCR (MSP) primers used to amplify promoter region of selected genes in BBM and in patients serum DNA.

Appendix A3.3

Gene	Primer	Primer sequence
CCDC8	F	AGG GCT ACA GTA GTT TGT TGG G
	R	GGC GGG CTT AGA GAT GAC T
BNC1	F	CAC TAC TTC ACA CCT GGG ATG
	R	GTT TGA GCT GTG TCT GTC TCT G
GALNT9	F	CAT CCC CAT GTC GAA CCC AG
	R	CAG ACA CTT GGA CTC AGG CAA G

Appendix A3.3: Primers designed for selected genes to carry out Reverse transcription PCR (RT-PCR).

Appendix B

Gene symbol	Accession	Gene name	% of tumours methylated	Function
ABCB1	NM_000927.4	*ATP-binding cassette sub-family B member 1*	80	Controls efflux of substances across plasma membranes, associated with multidrug resistance (Muggerud *et al.*, 2010)
AK5	NM_174858.2	*Adenylate kinase*	0	Involved in phosphoryl exchange (Solaroli *et al.*, 2009)
ALDH1A3	NM_000693.2	*Aldehyde dehydrogenase*	0	Retinal oxidation (Vasiliou *et al.*, 2000)
ANK3	NM_020987.3	*Ankyrin-3/G*	0	Regulate voltage gated sodium channel (Ferreira *et al.*, 2008)
ATM	NM_000051.3	*Ataxia telangiectasia mutated*	0	Key regulator in multiple signalling cascades (Rotman & Shiloh, 1998)
BNC1	NM_001717.3	*Basonuclin 1*	73	Zink finger transcription factor, regulator of EMT (Feuerborn *et al.*, 2014)
BOLL	NM_197970.2	*boule-like*	13	Associated with spermatogenesis (Westerveld *et al.*, 2005)
BVES	NM_007073.4	*Blood vessel epicardial substance*	64	Involved in inter-cellular interaction and cell adhesion. (Osler *et al.*, 2006)
CCDC8	NM_032040.4	*Coil coiled domain containing 8*	87	Mutated in patients with 3M syndrome (Hanson *et al.*, 2011). Loss is associated with genomic instability and aneuploidy (Yan *et al.*, 2014).
CD44	NM_000610.3	*CD44 molecule*	0	Main receptor for extracellular matrix component Hyaluronan (HA), associated in metastatic spread in various cancers (Gvozdenovic *et al.*, 2013)
CDKN1A	NM_000389.4	*Cyclin-Dependent*	7	Important component to regulate cell cycle

		Kinase Inhibitor 1A		(Shiozaki *et al.*, 2013)
CLDN1	NM_021101.4	*Claudin 1*	20	Intercellular adhesion molecule responsible for tight junction strand formation, its down/upregulation is associated with various cancers (Krämer *et al.*, 2000; Tsukita & Furuse, 2000; Fritzsche *et al.*, 2008; Myal *et al.*, 2010)
CLDN11	NM_005602.5	*Claudin 11*	0	Intercellular adhesion molecule responsible for tight junction strand formation (Tsukita & Furuse, 2000)
CLDN18	NM_016369.3	*Claudin 18*	100	Intercellular adhesion molecule responsible for tight junction strand formation (Tsukita & Furuse, 2000)
CLDN23	NM_194284.2	*Claudin 23*	0	Intercellular adhesion molecule responsible for tight junction strand formation, downregulated in various cancers (Tsukita & Furuse, 2000; Turksen & Troy, 2004)
CLDN3	NM_001306.3	*Claudin 3*	0	Intercellular adhesion molecule responsible for tight junction strand formation, its down/upregulation is associated with various cancers (Tsukita & Furuse, 2000; Turksen & Troy, 2004)
CLDN4	NM_001305.4	*Claudin 4*	0	Intercellular adhesion molecule responsible for tight junction strand formation, associated with various cancers and metastasis (Tsukita & Furuse, 2000; Turksen & Troy, 2004)
CLDN5	NM_001130861.1	*Claudin 5*	47	Intercellular adhesion molecule responsible for tight junction strand formation, involved in breast cancer metastasis, its loss affects blood brain barrier selectivity (Tsukita & Furuse, 2000; Nitta *et al.*, 2003; Escudero-

				Esparza *et al.*, 2012)
CLDN6	NM_021195.4	*Claudin 6*	55	Intercellular adhesion molecules responsible for tight junction strand formation, its epigenetic silencing is associated with migration and invasiveness of breast cancer (Tsukita & Furuse, 2000; Osanai *et al.*, 2007)
CLDN7	NM_001185022.1	*Claudin 7*	0	Intercellular adhesion molecule responsible for tight junction strand formation, its loss is associated with breast cancer progression (Tsukita & Furuse, 2000; Kominsky *et al.*, 2003)
CLDN9	NM_020982.3	*Claudin 9*	0	Intercellular adhesion molecule responsible for tight junction strand formation (Tsukita & Furuse, 2000)
CMTM8	NM_178868.3	*CKLF-like MARVEL transmembrane domain containing 8*	0	Negative regulator of EGM induced signalling, induces apoptosis (Jin *et al.*, 2007)
COL14A1	NM_021110.2	*Collagen, type XIV, alpha-1*	13	Interacts with extracellular matrix components, play roles in cell adhesion (Bauer *et al.*, 1997)
CSNK1A1	NM_001025105.2	*Casein kinase 1, alpha 1*	0	Serine/Threonine kinase plays role in cellular senescence (Sarasqueta *et al.*, 2013)
DAPK	NM_004938.3	*Death Associated protein kinase*	13	Serine/threonine kinase involved in apoptotic system (Pulling *et al.*, 2009) (Toyooka *et al.*, 2003)
DGKI	NM_004717.2	*Diacylglycerol kinase*	13	Activator of various signalling proteins (Regier *et al.*, 2005)
DLC1	NM_182643	*Deleted in Liver Cancer*	0	GTPase activating protein plays roles in signalling pathways (Wong *et al.*, 2003)
DSP	NM_004415.2	*Desmoplakin*	0	Inhibits Wnt/Beta catenin pathway, involved in cancers and metastasis (Davies *et al.*, 1999; Pang *et al.*, 2004; Chun & Hanahan, 2010; Yang

				et al., 2012)
FBN2	NM_001999.3	*fibrillin 2*	0	Associated with formation of microfibrils and elastic fibrillogenesis (Chaudhry *et al.*, 2001)
FBXL14	NM_152441.2	*F-box and leucine-rich repeat protein 14*	0	Involved with hypoxia and cancer development (Vinas-Castells *et al.*, 2010; Zheng *et al.*, 2012)
GALNT9	NM_001122636.1	*N-acetylgalactosaminyltransferase 9*	55	Catalyzes O-glycosylation (Shinya *et al.*, 2000; Berois *et al.*, 2013)
GATA5	NM_080473.4	*GATA binding protein 5*	0	Transcription factor, activates anti-tumour genes and act as TSG (Akiyama *et al.*, 2003)
GREM1	NM_013372.6	*Gremlin1*	7	Plays role in tissue modelling and angiogenesis (Mulvihill *et al.*, 2012)
HK2	NM_000189.4	*Hexokinase*	0.	First rate limiting enzyme of glycolysis, involved in pancreatic carcinogenesis Dong *et al.* (2011) (Fang *et al.*, 2012)
HOXB13	NM_006361.5	*Homeobox B13*	80	TSG for prostate cancer, inhibits androgen mediated signalling (Fidler *et al.*, 2010)
HOXD10	NM_002148.3	*Homeobox D10*	55	Maintain epithelial cell plasticity and contributes to stability of extracellular matrix (Carrio *et al.*, 2005)
HOXD3	NM_006898.4	*HomeoboxD3*	100	Proangiogenic transcription factor (Chen *et al.*, 2004)
ICAM5	NM_003259.3	*intercellular adhesion molecule 5*	0	Involved in dendritic outgrowth, associated with prostate cancer (Griffith & Swartz, 2006)
IGFBP3	NM_000598.4	*Insulin-like growth factor-binding protein 3*	0	Supresses tumour growth and induce apoptosis (Regel *et al.*, 2012)
KLHL35	NM_001039548.2	*Kelch –like 35*	0	Associated with Mendalian diseases and cancers (Dhanoa *et al.*, 2013)
KRT18	NM_000224.2	*Keratin 18*	0	Component of intermediate filament

				in epithelial cells contributing to cell to cell adhesions, involved in apoptosis and is associated with invasiveness of breast cancer (Magin *et al.*, 2007; Bragulla & Homberger, 2009; Iyer *et al.*, 2013) (Oshima, 2002)
KRT19	NM_002276.4	*Keratin 19*	0	Component of intermediate filament in epithelial cells contributing to cell to cell adhesion (Magin *et al.*, 2007; Bragulla & Homberger, 2009)
KRT28	NM_181535.3	*Keratin 28*	0	Component of intermediate filament in epithelial cells contributing to cell to cell adhesions (Magin *et al.*, 2007; Bragulla & Homberger, 2009)
KRT7	NM_005556.3	*Keratin 7*	0	Component of intermediate filament in epithelial cells contributing to cell to cell adhesions (Magin *et al.*, 2007; Bragulla & Homberger, 2009)
KRT72	NM_080747.2	*Keratin 72*	42	Component of intermediate filament in epithelial cells in contributing to cell to cell adhesion (Magin *et al.*, 2007; Bragulla & Homberger, 2009)
KRT81	NM_002281.3	*Keratin 81*	0	Component of intermediate filament in epithelial cells contributing to cell to cell adhesions (Magin *et al.*, 2007; Bragulla & Homberger, 2009)
KRT83	NM_002282.3	*Keratin 83*	84	Component of intermediate filament, contributes to cell to cell adhesion (Magin *et al.*, 2007; Bragulla & Homberger, 2009)
KRT85	NM_002283.3	*Keratin 85*	100	Component of intermediate filament in epithelial cells contributing to cell-cell adhesion (Magin *et al.*, 2007; Bragulla & Homberger, 2009;

				Shimomura *et al.*, 2010)
KRT86	NM_002284.3	*Keratin 86*	0	Component of intermediate filament in epithelial cells contributing to cell to cell adhesions (Magin *et al.*, 2007; Bragulla & Homberger, 2009)
MIR124-1	NR_029668.1	*microRNA 124-1*	*36*	miRNA with tumour suppressor activity, epigenetically deregulated in various cancers (Wilting *et al.*, 2010; Wong *et al.*, 2011; Shi *et al.*, 2013)
MIR127	NR_029696.1	*microRNA 127*	100	Regulator of cell proliferation and senescence (Chen *et al.*, 2013)
MIR23B	NR_029664.1	*microRNA 23b*	92	Involved in cytoskeleton modelling, motility and metastasis (Majid *et al.*, 2012; Zaman *et al.*, 2012; Jin *et al.*, 2013; Pellegrino *et al.*, 2013)
MIR34a	NR_029610.1	*microRNA 34*	10	Transcriptional target of p53 associated with various cancers, apoptosis and metastasis (Rokhlin *et al.*, 2008; Yamakuchi *et al.*, 2008; Hermeking, 2010; Roy *et al.*, 2012)
MIR34b	NR_029839.1	*microRNA 34b*	30	Tumour suppressor miRNA, associated with p53 regulation, cancer and apoptosis (Hermeking, 2010)
MIR433	NR_029966.1	*microRNA 433*	100	Deregulated in gastric cancer, regulator of cell migration and drug response (Luo *et al.*, 2009; Symmans, 2010)
MMP2	NM_004530.4	*Matrix metallopeptidase 2*	0	A type of extracellular enzymes that digests extracellular membranes and promotes cell migration/ and promotes tumour development invasion of (Somiari *et al.*, 2006)
MST1R	NM_002447.2	*macrophage stimulating 1 receptor*	78	Involved in intracellular signalling cascades leading to cellular growth,

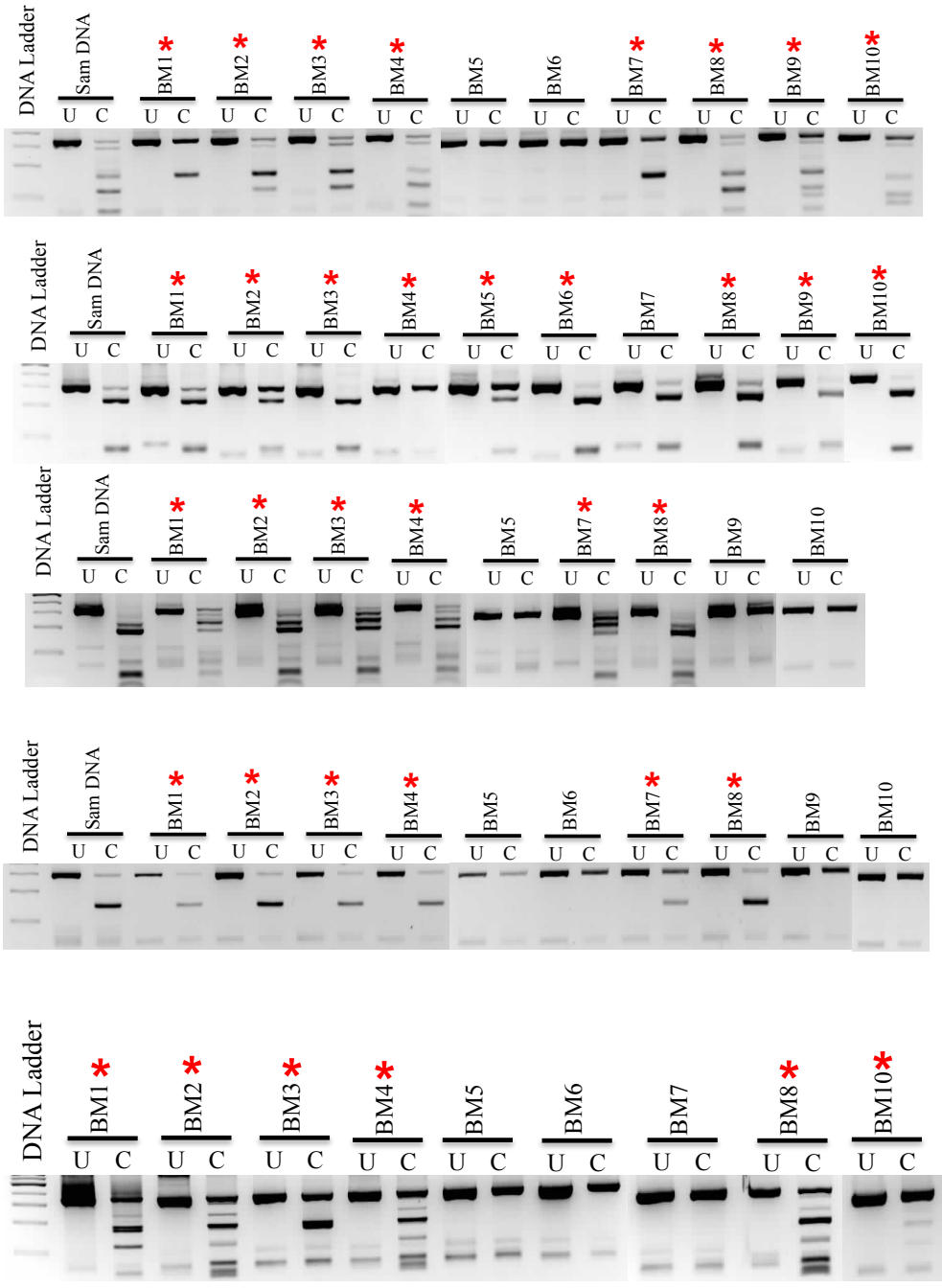
				motility and invasion (Wagh *et al.*, 2008)
NRCAM	NM_001037132.2	*Neuronal cell adhesion molecule*	0	Functions as neural-glial cell adhesion molecule, plays role in neural proliferation, proliferation and signalling pathways (Sakurai, 2012)
OCLN	NM_002538.3	*Occludin*	0	Tight junction associated integral protein, (Martin *et al.*, 2010)
PBRM1	NM_018313.4	*polybromo 1*	0	Regulates embryonic development associated with renal cell carcinoma (Varela *et al.*, 2011b)
PCDH8	NM_002590.3	*Procadhern 8*	73	Helps in cell to cell adhesion (Sabine *et al.*, 1998)
PENK	NM_006211.3	*Proenkephalin*	80	Promotes RNA splicing in osteoblasts and neural cells, plays role in bone development (Rosen *et al.*, 2013a)
PNN	NM_002687.3	*pinin, desmosome associated protein*	0	Associated with linking intermediate filaments to desmosome in epithelial cells (Ouyang & Sugrue, 1996)
PTEN	NM_000314	*Phosphatase and tensin homolog*	0	Modulates cell signalling, growth, migration and apoptosis, regulates Pi3K pathway (Wu *et al.*, 1998; Yamada & Araki, 2001)
PYCARD	NM_013258.4	*Apoptosis-associated speck-like protein containing a CARD*	0	Pro-apoptotic TSG (Siraj *et al.*, 2011)
QPCT	NM_012413.3	*Glutaminyl-peptide cyclotransferase*	0	CCL2 signalling, may be involved in thyroid carcinoma (Kehlen *et al.*, 2012)
RBP1	NM_002899.3	*retinol binding protein 1*	0	Involved in retinol transport and metabolism, associated with ovarian cancer (Stephens *et al.*, 2012)
SDHD	NM_001276503.1	*Succinate dehydrogenase complex, subunit D*	0	Encodes respiratory chain protein, associated with paraganglioma (Gimenez-Roqueplo *et*

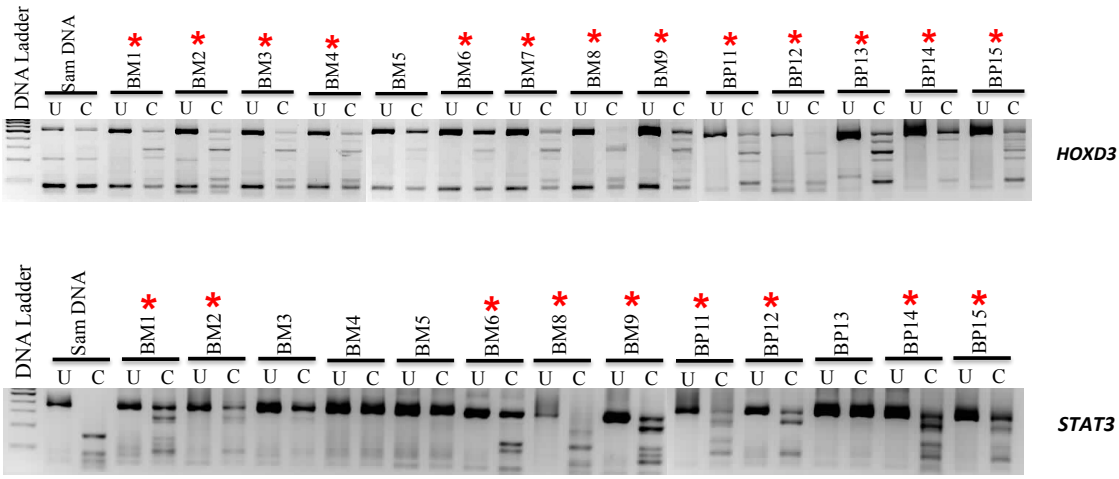
				al., 2001)
SFRP2	NM_003013.2	*Secreted frizzled-related protein 2*	47	Regulator of canonical Wnt pathway (Veeck *et al.*, 2008)
STAT3	NM_139276.2	*Signal transducer and activator of transcription 3*	67	Involved in embryonic stem cell regulation, somatic cell growth (KIYOSHI *et al.*, 1997; Hitoshi *et al.*, 1998; Akira, 2000)
SULF2	NM_018837.3	*sulfatase 2*	0	Modulates various signalling proteins and inhibit tumour growth (Chau *et al.*, 2009)
TFAP2A	NM_003220.2	*Transcription factor AP-2 alpha*	0	Required for neural crest induction (Li & Cornell, 2007)
TJP1	NM_003257.3	*tight junction protein 1*	0	Plays role in tight junction organization and assembly (D'Atri *et al.*, 2002)
TMEFF2	NM_016192.2	*Transmembrane protein with EGF-like and two follistatin-like domains 2*	0	Transmembrane protein, suppresses prostate tumour (Gery & Koeffler, 2003)
TNFRSF10C	NM_003841.3	*Tumour Necrosis Factor receptor superfamily 10C*	13	Cell surface protein, modulates multiple biological networks (Degli-Esposti *et al.*, 1997)
TNFRSF10D	NM_003840.4	*Tumour Necrosis Factor receptor superfamily 10 D*	60	Member of TNF(Tumour Necrosis Factor) receptor superfamily, promotes apoptosis in cancer cells (Hill *et al.*, 2011)
TSC1	NM_000368.4	*Tuberous Sclerosis 1*	0	May be involved in tumour suppression (Miloloza *et al.*, 2000)
TSPAN13	NM_014399.3	*Tetraspanin 13*	0	Transmembrane protein, inhibits cellular growth and invasion (Arencibia *et al.*, 2009)
TSPAN4	NM_001025237.1	*Tetraspanin 4*	0	Cellular growth, adhesion and differentiation (Todd *et al.*, 1998b)
UCHL1	NM_004181.4	*Ubiquitin carboxyl-terminal esterase L1*	0	Neuron specific ubiquitinating enzyme, associated with early onset of progressive neurodegeneration(Bilguvar *et al.*, 2012)
WIF1	NM_007191.4	*Wnt*	53	Inhibitor of Wnt-

		inhibitory factor-1 gene		signalling (Ai *et al.*, 2006; Veeck *et al.*, 2009)
ZNF808	NM_001039886.3	*Zink finger protein 808*	0	Zink finger protein, may be involved in transcriptional regulation (Lancet *et al.*, 2013)

Appendix B1: A list of all genes screened for methylation status in BBM

Appendix B2:





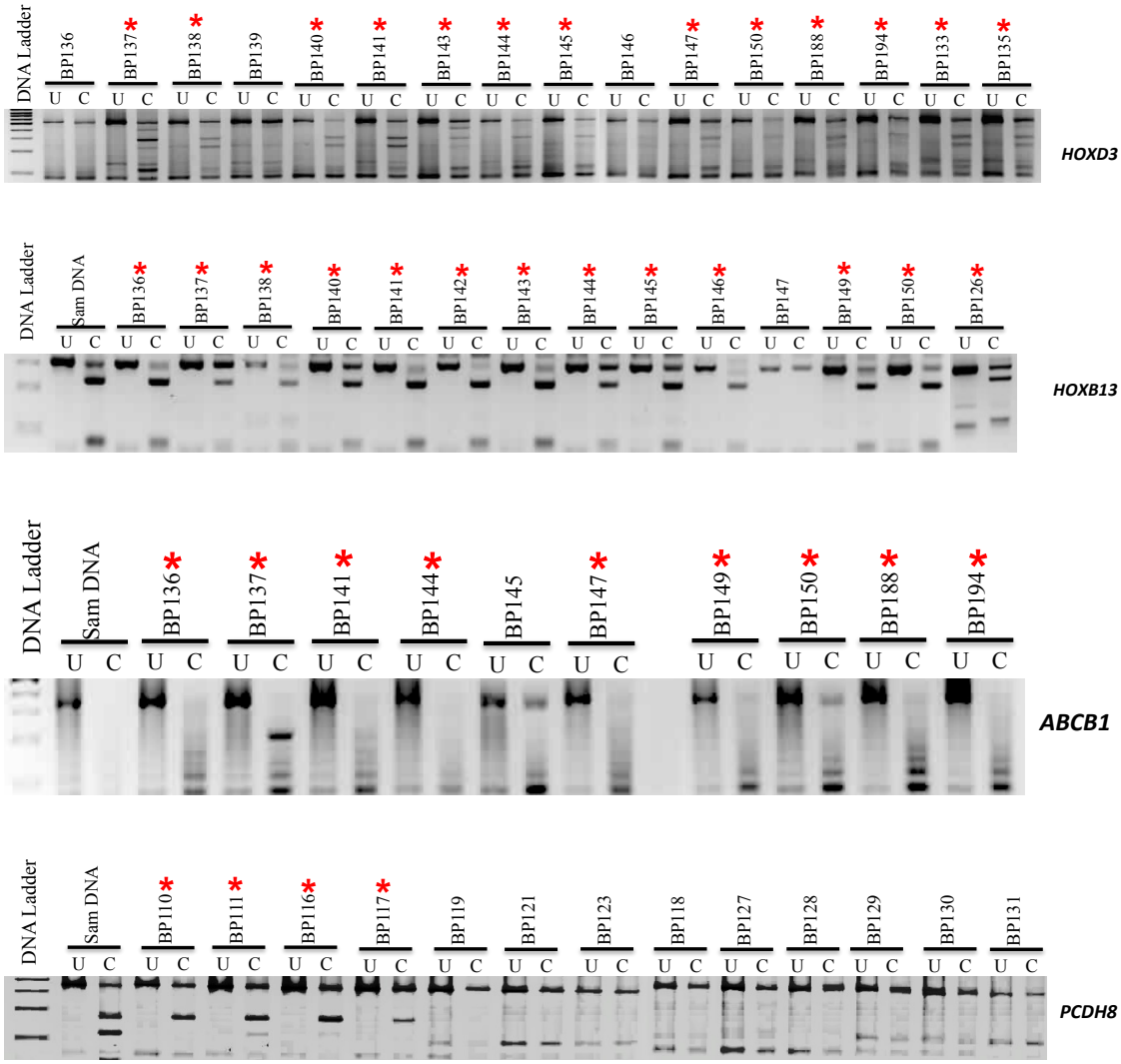
Appendix B2: Gel electrophoresis images of some additional genes, which were frequently methylated in BBM samples.

Appendix B4:

Estrogen receptor status	Progesterone receptor status	HER2 status	Date of surgery	Age at diagnosis	Date of investigation for metastasis	Metastasis	Bone metastasis	Lung metastasis	Cutaneous metastasis	Multiples metastases	Hepatic metastasis	Brain metastasis	Lymph node
P	P	N	1978	37	1981	1	1	0	0	0	1	0	
N	N	N	1984	59	1999	0	0	0	0	0	0	0	0
N	N	N	1984	71	2001	0	0	0	0	0	0	0	0
P	P	N	1985	79	1991	0	0	0	0	0	0	0	0
P	N	N	1985	73	1988	1	0	0	0	0	1	0	
N	N	P	1986	83	1992	0	0	0	0	0	0	0	0
P	P	N	1991	72	1998	1	1	1	0	1	1	0	
P	P	P	1987	50	2013	0	0	0	0	0	0	0	0
N	N	P	1987	63	2008	0	0	0	0	0	0	0	0
P	P	N	1989	64	2004	1	1	0	0	0	0	0	
P	P	N	1987	66	1995	0	0	0	0	0	0	0	0
P	P	P	1987	43	2006	0	0	0	0	0	0	0	0
N	N	P	1987	75	2000	0	0	0	0	0	0	0	0
P	P	N	1990	81	2000	1	0	0	0	0	1	0	
P	P	N	11981	68	1989	1	0	0	1	0	0	0	
P	P	N	1991	78	1999	0	0	0	0	0	0	0	0
P	P	N	1991	80	2002	0	0	0	0	0	0	0	0
P	N	N	1992	70	1999	0	0	0	0	0	0	0	0
P	P	N	1992	59	1997	0	0	0	0	0	0	0	0
P	P	N	1984	62	1998	0	0	0	0	0	0	0	0

Appendix B4: Clinical information of unrelated primary tumours

Appendix B5:



Appendix B5: Gel electrophoresis images of some additional genes frequently methylated in unrelated primary breast tumour

Appendix B7

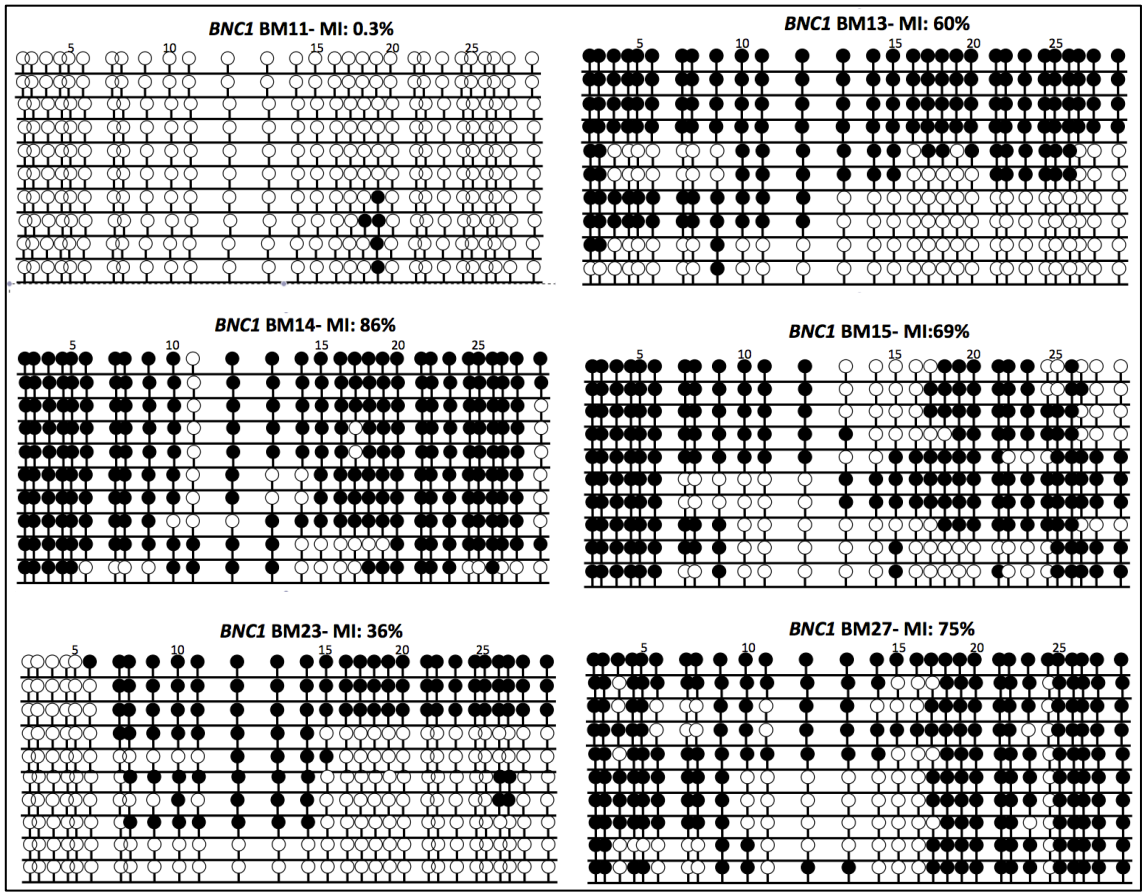
Tumour	Estrogen receptor status	Progesterone receptor status	HER2 status	Lymph/Vascular Invasion	Duration between primary breast tumours and BBM surgery
BP1	Positive	Negative	Negative	Vascular invasion noted	5 years
BP2	Positive	Positive	Negative	NA	10 years
BP3	Positive	Positive	Negative,	Lymphovascular invasion	2 years
BP5	Positive	Negative	Negative	NA	10 years
BP8	NA	NA	NA	NA	2 years
BP10	Negative	Negative	NA	NA	3 years
BP11	NA	NA	1+ (negative)	NA	6 years
BP12	Positive	Positive	1+ (negative)	Lymphovascular invasion	5 years
BP13	Negative	Negative	Negative	NA	3 years
BP14	Positive	Negative	Negative	NA	4 years
BP15	NA	NA	NA	Lymphovascular invasion	2 years

Appendix B7: Clinical information on matched primaries that have metastasised to the brain.

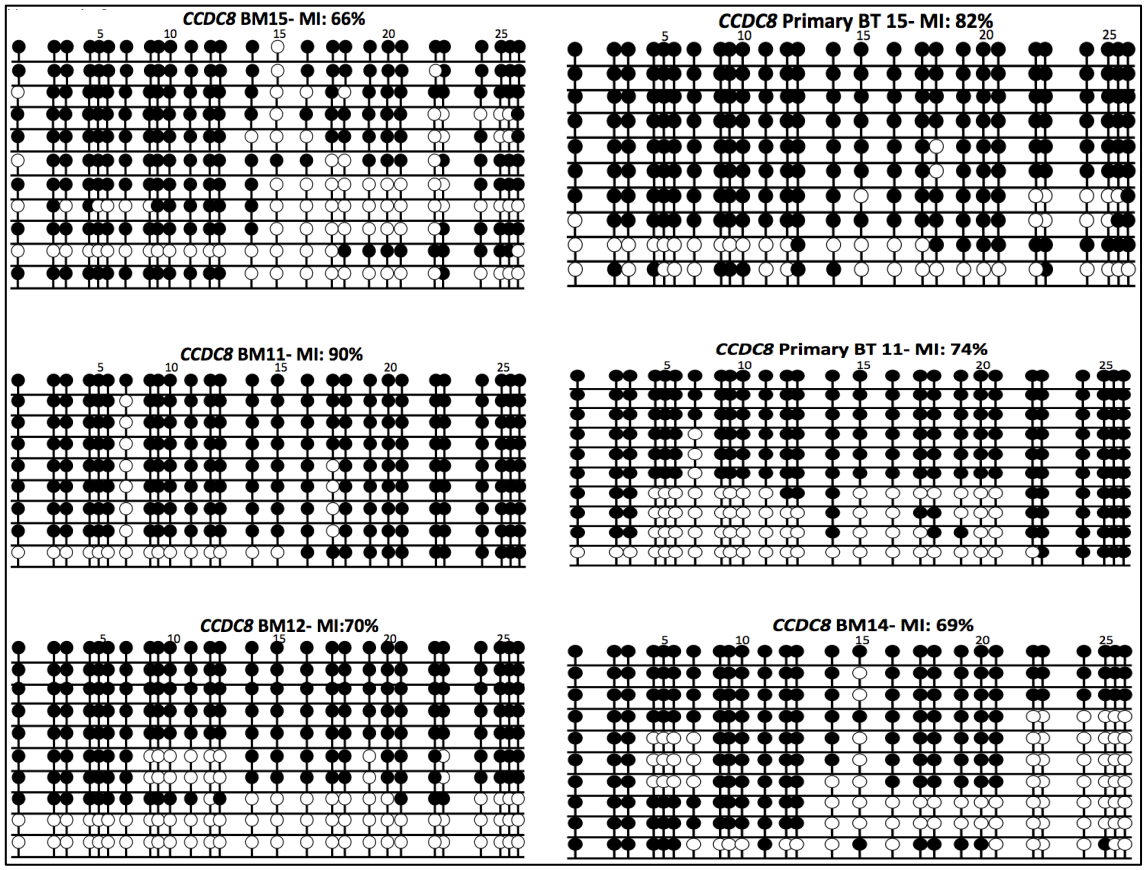
Appendix C

Appendix C1: Clone sequencing images

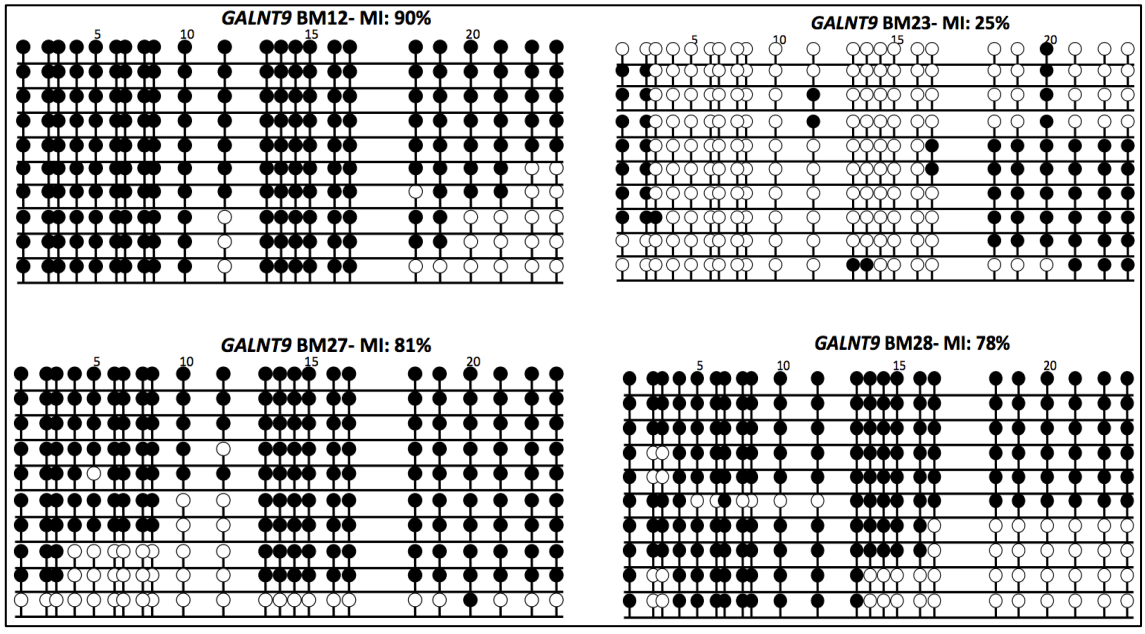
Appendix C1. 1



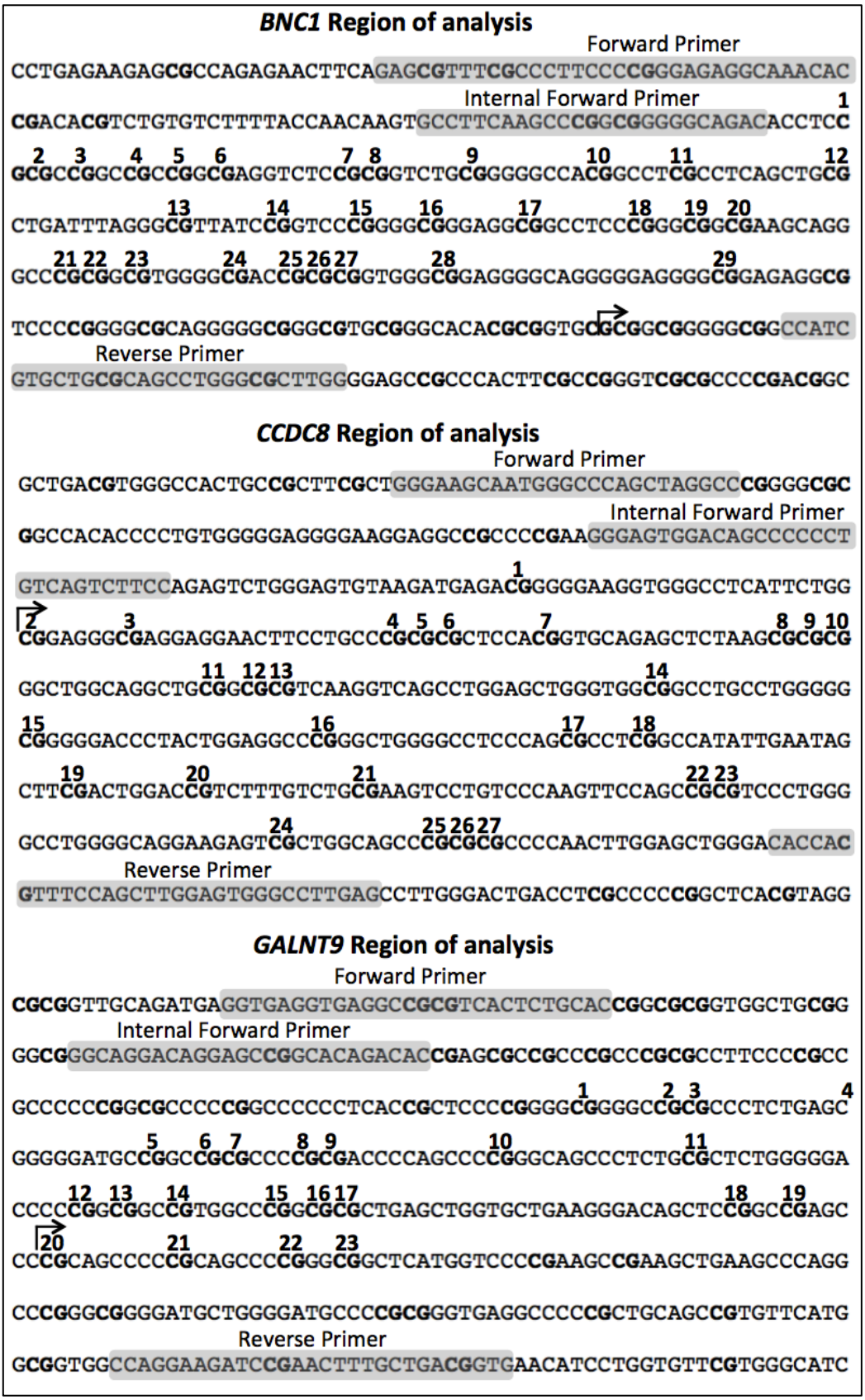
Appendix C1.1: Clone sequencing: BNC1 methylation in individual alleles



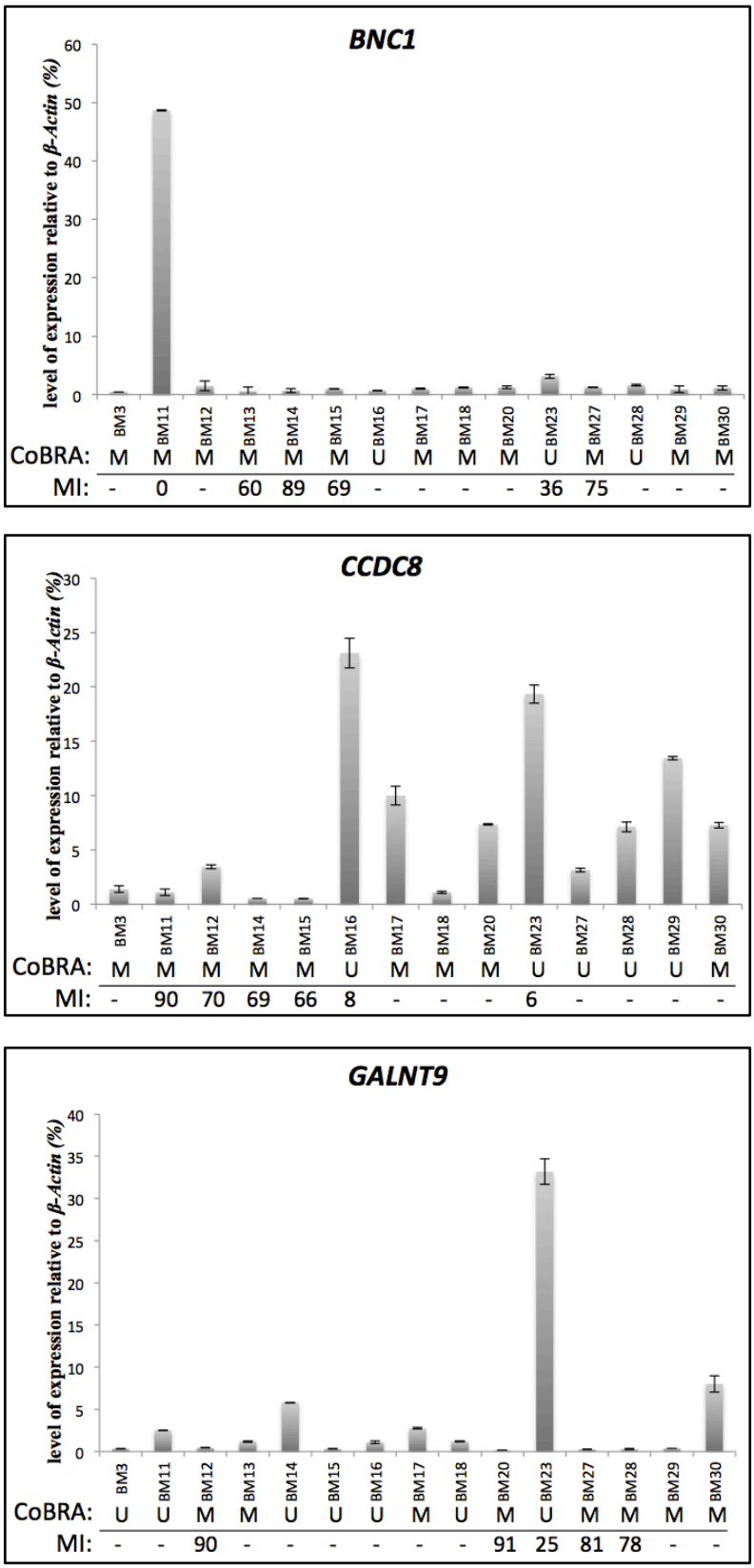
C1.2 Clone sequencing of *CCDC8* methylation in individual alleles



C1.3 Clone sequencing: *GALNT9* methylation in individual alleles



C1.4 Region of analysis of base pair resolution of methylation in individual alleles

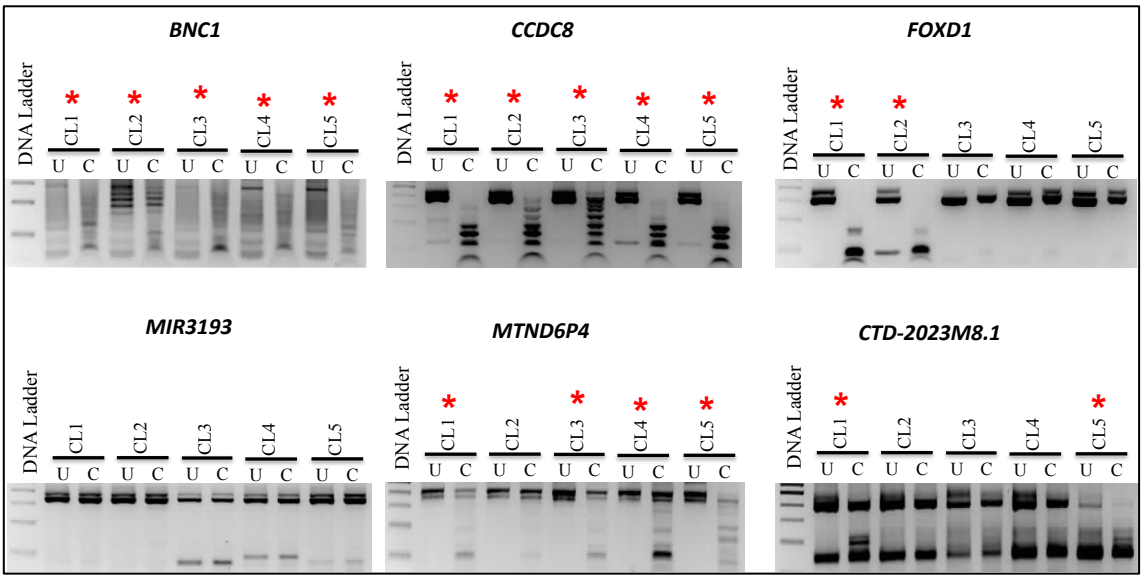


Appendix C2 Quantitative expression of three genes *BNC1*, *CCDC8* and *GALNT9* in BBM samples

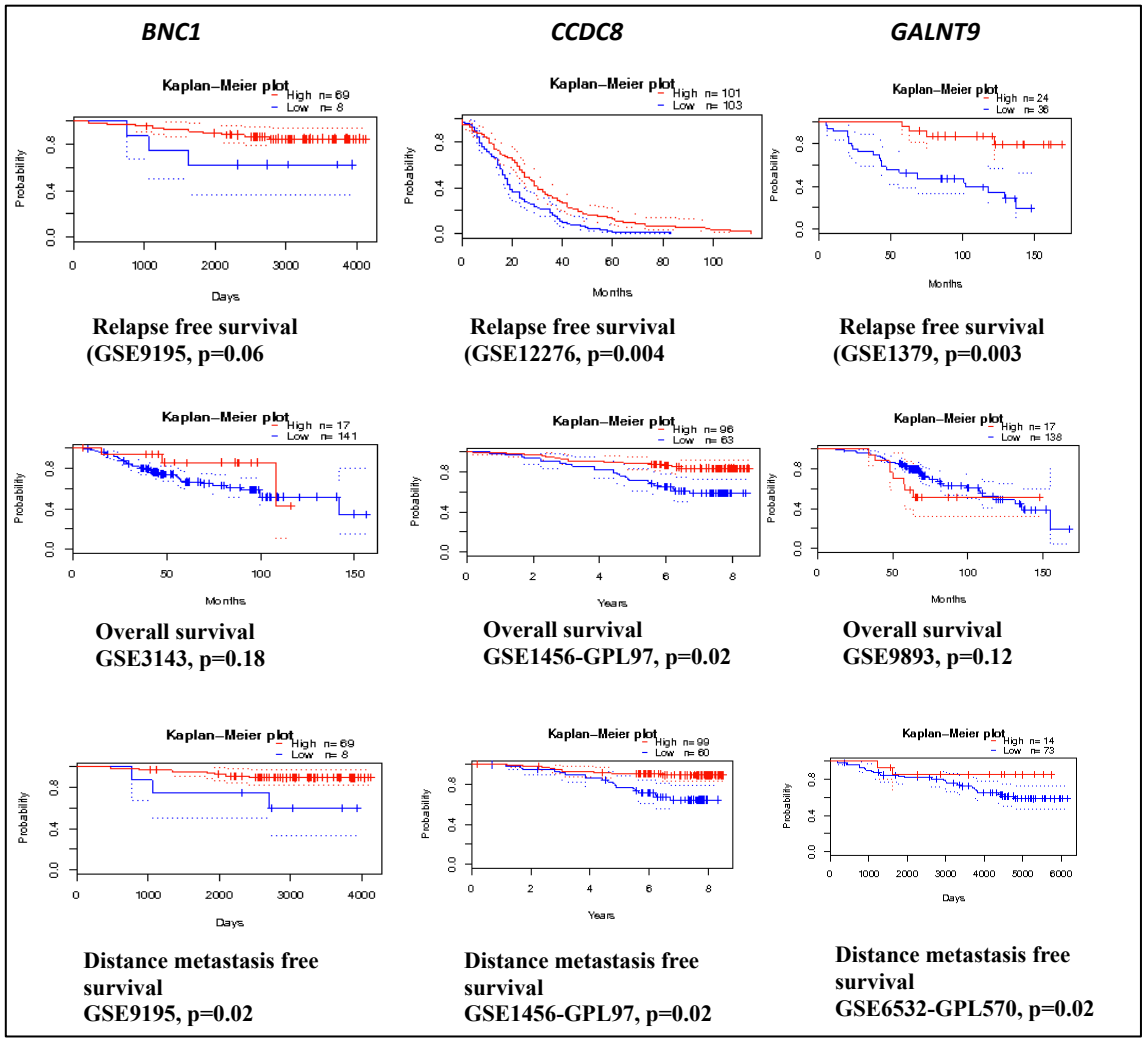
Metastatic and Non metastatic primary breast tumours from the TCGA used for the clinical information			
BRCA METASTATIC tumours		BRCA NON METASTATIC tumours	
Tumour Code	Tumour barcode	Tumour Code	Tumour barcode
M1	TCGA-A2-A3XS-01A-11D-A230-05	NM1	TCGA-D8-A1X6-01A-11D-A14N-05
M2	TCGA-EW-A1P8-01A-11D-A145-05	NM2	TCGA-D8-A1XJ-01A-11D-A14N-05
M3	TCGA-EW-A1P1-01A-31D-A14H-05	NM3	TCGA-D8-A1XM-01A-21D-A14N-05
M4	TCGA-AR-A2LH-01A-31D-A18O-05	NM4	TCGA-D8-A1XO-01A-11D-A14N-05
M5	TCGA-GM-A2D9-01A-11D-A18O-05	NM5	TCGA-D8-A1XQ-01A-11D-A14N-05
M6	TCGA-GM-A2DA-01A-11D-A18O-05	NM6	TCGA-D8-A1XR-01A-11D-A14N-05
M7	TCGA-AC-A2FM-01A-11D-A19Z-05	NM7	TCGA-D8-A1XS-01A-11D-A14N-05
M8	TCGA-EW-A1P0-01A-11D-A145-05	NM8	TCGA-D8-A1XT-01A-11D-A14N-05
M9	TCGA-BH-A1FH-01A-12D-A13K-05	NM9	TCGA-D8-A1XU-01A-11D-A14N-05
M10	TCGA-A2-A0SW-01A-11D-A10P-05	NM10	TCGA-D8-A1XV-01A-11D-A14N-05
M11	TCGA-A2-A0T2-01A-11D-A10P-05	NM11	TCGA-D8-A1XW-01A-11D-A14N-05
M12	TCGA-AR-A0TZ-01A-12D-A10P-05	NM12	TCGA-D8-A1XY-01A-11D-A14N-05
M13	TCGA-AR-A0U2-01A-11D-A10A-05	NM13	TCGA-D8-A1XZ-01A-11D-A14N-05
M14	TCGA-A2-A0SV-01A-11D-A10P-05	NM14	TCGA-D8-A1Y0-01A-11D-A14N-05

Appendix C3: Metastatic and non-metastatic primary tumours from the TCGA used for clinical information.

Appendix D



Appendix D1: cobra methylation for cell lines. CL1: MCF7, CL2: T47D, CL3: MDA-MB231, CL4: BT549, CL5: ZR75, U: Uncut (control), C: Cut by restriction enzyme (BstU1/TaqI), *: methylated cell line



Appendix D2: Survival curves for three genes *BNC1*, *CCDC8* and *GALNT9* in different data sets

Appendix E

Normal breast tissues and primary breast tumours (Invasive breast carcinoma BRCA) used for comparing with BBM data to screen candidates for BBM	
Normal breast tissues from the TCGA	**Primary breast tissues from the TCGA**
Tumour barcode	**Tumour barcode**
TCGA-AC-A23H-11A-12D-A161-05	TCGA-D8-A1X6-01A-11D-A14N-05
TCGA-BH-A204-11A-53D-A161-05	TCGA-D8-A1XJ-01A-11D-A14N-05
TCGA-BH-A209-11A-42D-A161-05	TCGA-D8-A1XM-01A-21D-A14N-05
TCGA-E9-A1RB-11A-33D-A161-05	TCGA-D8-A1XO-01A-11D-A14N-05
TCGA-E2-A1LI-11A-23D-A161-05	TCGA-D8-A1XQ-01A-11D-A14N-05
TCGA-E2-A1LS-11A-32D-A161-05	TCGA-D8-A1XR-01A-11D-A14N-05
TCGA-E9-A1RC-11A-33D-A161-05	TCGA-D8-A1XS-01A-11D-A14N-05
TCGA-E9-A1RD-11A-33D-A161-05	TCGA-D8-A1XT-01A-11D-A14N-05
TCGA-BH-A208-11A-51D-A161-05	TCGA-D8-A1XU-01A-11D-A14N-05
TCGA-E9-A1RF-11A-32D-A161-05	TCGA-D8-A1XV-01A-11D-A14N-05
TCGA-AC-A23H-11A-12D-A161-05	TCGA-D8-A1XW-01A-11D-A14N-05
TCGA-AC-A23H-11A-12D-A161-05	TCGA-D8-A1XY-01A-11D-A14N-05
	TCGA-D8-A1XZ-01A-11D-A14N-05
	TCGA-D8-A1Y0-01A-11D-A14N-05

E1: TCGA Barcode identifiers of normal tissues and primary tumours used in analyses of 450K methylation data in BBM

Gene Symbol	Entrez ID	Chromosome	Mean (β value) BBM	Mean (β value) primary tumours	Difference	Promoter
HSPB9	94086	chr17	2.029788864	0.0868493	2.130872037	TRUE
DDX52	11056	chr17	1.525909178	0.417176197	1.118207247	TRUE
EDARADD	128178	chr1	0.422072021	0.726856306	1.192213701	TRUE
RNF8	157360	chr6	0.822653313	0.524739581	0.297913731	TRUE
SOX5	281485	chr12	0.837062213	0.556376304	0.28068591	TRUE
RBM23	311885	chr14	0.818144254	0.538907402	0.279236851	TRUE
LMX1B	224781	chr9	0.826886808	0.54627828	0.280608528	TRUE
OTX1	5013	chr2	-2.052294171	-0.963869274	-1.459176341	TRUE
NDUFS6	4726	chr5	-2.363352982	-1.068901118	-1.715868344	TRUE
MEIS2	4212	chr15	-2.262082477	-1.023659256	-1.504283882	TRUE
NR2F1	7025	chr5	-1.67850661	-0.204925415	-1.473581195	TRUE
MIR1178	100302274	chr12	1.752820322	0.458007976	1.403570235	TRUE
SLITRK5	26050	chr13	-2.779276933	-1.879359928	-1.057982323	TRUE
CSRP2	1466	chr12	-2.542994118	-1.651557161	-1.194480426	TRUE
LOC154872	154872	chr7	-0.076256492	-1.326517102	1.461613651	TRUE
RPS6KA2	6196	chr6	-2.629307437	-1.685474923	-1.111214542	TRUE
BCL9	607	chr1	-1.710535721	-0.699173354	-1.06067673	TRUE
KCNA2	3737	chr1	-1.87471767	-0.88162023	-1.035807156	TRUE
PANK1	53354	chr10	-2.660896875	-1.6144809	-1.046415975	TRUE
PIP5K1A	8394	chr1	-1.924748424	-1.088080394	-1.167612619	TRUE
FLJ22447	400221	chr14	-2.16613857	-1.09760256	-1.06853601	TRUE
ALDHL1	89880	chr3	0.615892326	0.255782622	0.360109704	TRUE
SCAND3	114821	chr6	2.052172389	0.801862888	1.250309501	FALSE
UNKL	64718	chr16	-1.066310228	0.164519429	-1.709470409	FALSE
PRDM2	7799	chr1	-2.241456505	-1.140906534	-1.100549971	FALSE
MIR5580	100847076	chr14	-2.21502181	-1.096813745	-1.425877581	FALSE
PAX9	5083	chr14	-2.548073287	-1.39339006	-1.568435629	FALSE
ZNF331	55422	chr19	-2.114654599	-1.142370508	-1.034999342	FALSE
LOC388942	388942	chr2	-2.654429853	-1.753323924	-1.079412215	FALSE
MNX1	3110	chr7	-0.917992399	0.52934354	-1.829976908	FALSE
RNF220	55182	chr1	-2.518476022	-1.043731152	-1.669193455	FALSE
SCD	6319	chr10	-1.669775868	-0.698550507	-1.09810695	FALSE
GNG4	2786	chr1	-1.504077952	-0.478997596	-1.097389678	FALSE
GATA3	2625	chr10	2.813936505	1.576045434	1.237891071	FALSE
PARG	8505	chr10	-1.553188052	-0.526076947	-1.027111105	FALSE
TLR3	7098	chr4	2.158430487	1.183355862	1.031898906	FALSE
FGF9	2254	chr13	-2.57297787	-1.450151353	-1.122826518	FALSE
UNQ6975	400952	chr2	-2.136015239	-1.002047522	-1.133967718	FALSE
NFIB	4781	chr9	-2.778820565	-1.465912041	-1.390347762	FALSE
ARID3A	1820	chr19	-2.097062072	-1.046837775	-1.093688548	FALSE
OSR2	116039	chr8	-2.764144155	-1.371592925	-1.44085651	FALSE
MECOM	2122	chr3	-2.497820446	-1.461679264	-1.278761723	FALSE
EFNA2	1943	chr19	1.977663492	2.995766323	-1.018102831	FALSE
KCTD8	386617	chr4	-0.871676426	0.157637684	-1.02931411	FALSE
HHIP	64399	chr4	-2.97240584	-1.918370607	-1.209099267	FALSE
PYY	5697	chr17	-2.96844562	-2.094559227	-1.059136332	FALSE
C1orf216	127703	chr1	-2.60973537	-1.600226301	-1.009509069	FALSE
SRSF7	6432	chr2	-0.321410898	0.937379935	-1.258790833	FALSE
SLC35F3	148641	chr1	1.990232239	0.930531034	1.059701205	FALSE
OR10AD1	121275	chr12	-1.157595374	0.32176395	-1.54475007	FALSE
DNAJB6	10049	chr7	-0.643603812	0.37155827	-1.106530684	FALSE
BNIP3	664	chr10	-1.514194422	-0.093777042	-1.420417379	FALSE
LINC00476	100128782	chr9	2.224419525	1.226652806	1.029069023	FALSE
MRGPRX1	259249	chr11	1.71317044	0.685085961	1.028084478	FALSE
TRABD	80305	chr22	-0.928343052	-2.14998339	1.221640338	FALSE
MARCKS	4082	chr6	-2.81590301	-1.574430944	-1.347103793	FALSE
SLC12A9	56996	chr7	-1.028879937	-0.086277383	-1.008520316	FALSE
ARMC2	84071	chr6	2.843974028	1.841212414	1.002761614	FALSE
CCDC80	151887	chr3	2.198874929	1.065549344	1.169874992	FALSE

Appendix E2: list of probes generated, which are differentially methylated in BBM compared in primary breast tumours from TCGA (based on mean β values of all tumours)

Probes	Ensembl Gene ID	Gene	Chomosome
cg25066665	ENSG00000163125	RPRD2	*chr1*
cg16736018	ENSG00000237588	RP11-66D17.3	*chr1*
cg13052638	ENSG00000261000	*Unknown*	*chr1*
cg26362491	ENSG00000229367	HMGN2P19	*chr1*
cg00659878	ENSG00000229367	HMGN2P19	*chr1*
cg24339574	ENSG00000224939	*LINC00184*	*chr1*
cg12720965	ENSG00000232192	*Unknown*	*chr1*
cg19749001	ENSG00000264010	*MIR4429*	*chr2*
cg25785303	ENSG00000264010	*MIR4429*	*chr2*
cg26563141	ENSG00000229604	*MT−ATP8*	*chr2*
cg13231117	ENSG00000229689	*AC009237.8*	*chr2*
cg01882471	ENSG00000239795	AC109826.2	*chr2*
cg27612889	ENSG00000138386	NAB1	*chr2*
cg17974460	ENSG00000222972	*Unknown*	*chr2*
cg14731570	ENSG00000261829	*Unknown*	*chr2*
cg01070987	ENSG00000206199	*ANKUB1*	*chr3*
cg10636490	ENSG00000206199	*ANKUB1*	*chr3*
cg09406615	ENSG00000114200	*BCHE*	*chr3*
cg12494166	*ENSG00000251129*	RP11-734I18.1	chr4
cg15885430	ENSG00000250020	RP11-811I15.1	chr5
cg03512172	ENSG00000251532	*Unknown*	*chr5*
cg11599887	ENSG00000251532	*Unknown*	*chr5*
cg27039593	ENSG00000251532	*Unknown*	*chr5*
cg08808615	ENSG00000251532	*Unknown*	*chr5*
cg18634758	ENSG00000251532	*Unknown*	*chr5*
cg16292885	ENSG00000251532	*Unknown*	*chr5*
cg18262197	ENSG00000251532	*Unknown*	*chr5*
cg18895088	ENSG00000251532	*Unknown*	*chr5*
cg21129181	ENSG00000251532	*Unknown*	*chr5*
cg21771528	ENSG00000248693	*CTD-2023M8.1*	*chr5*
cg21806580	ENSG00000248693	*CTD-2023M8.1*	*chr5*
cg22422937	ENSG00000251675	*Unknown*	*chr5*
cg23311108	ENSG00000185641	*Unknown*	*chr5*
cg12949141	ENSG00000249119	*MTND6P4*	*chr5*
cg06793849	ENSG00000253925	*Unknown*	*chr5*
cg11061434	ENSG00000253925	*Unknown*	*chr5*
cg10647925	ENSG00000253925	*Unknown*	*chr5*
cg22871227	SG00000249031/SUMO2P6		*chr5*
cg14429919	ENSG00000229282	*Unknown*	*ch6*
cg00660009	ENSG00000229282	*Unknown*	*ch6*
cg24395504	ENSG00000220181	*Unknown*	*ch6*
cg01196858	ENSG00000227014	*Unknown*	*chr7*
cg06133110	ENSG00000227014	*Unknown*	*chr7*
cg14399183	ENSG00000235865	*GSN−AS1*	*chr9*
cg12867448	ENSG00000136811	*ODF2*	*chr9*
cg09337049	ENSG00000242853	*Unknown*	*chr10*
cg24232869	ENSG00000226138	*SENP1/SUMO1*	*chr12*
cg16264616	ENSG00000250133	*HOXC−AS2*	*chr12*
cg19794507	ENSG00000250133	*HOXC−AS2*	*chr12*
cg09941406	ENSG00000250133	*HOXC−AS2*	*chr12*
cg20340866	ENSG00000250133	*HOXC−AS2*	*chr12*
cg06208615	ENSG00000257342	*Unknown*	*chr12*
cg11899507	ENSG00000200135	*Unknown*	*chr12*
cg21532408	ENSG00000134864	A2LD1	chr13
cg10256726	ENSG00000151327	*FAM177A1*	*chr14*
cg19664425	ENSG00000258377	*Unknown*	*chr14*
cg13704629	ENSG00000176043	*Unknown*	*chr14*
cg24043861	ENSG00000176043	*Unknown*	*chr14*
cg13158272	ENSG00000156030	*C14orf43*	*chr14*
cg10356657	ENSG00000258749	Unknown	*chr14*
cg06897120	ENSG00000261612	*SUB1P3*	*chr16*
cg20438687	ENSG00000230201	*ATP6V0CP1*	*chr17*
cg12608565	*ENSG00000205710*	*C17orf107*	*chr17*
cg05529816	*ENSG00000205710*	*C17orf107*	*chr17*
cg14275842	*ENSG00000205710*	*C17orf107*	*chr17*
cg07834574	*ENSG00000205710*	*C17orf107*	*chr17*
cg09036188	*ENSG00000205710*	*C17orf107*	*chr17*
cg14482741	*ENSG00000205710*	*C17orf107*	*chr17*
cg20814095	*ENSG00000205710*	*C17orf107*	*chr17*
cg21071097	*ENSG00000205710*	*C17orf107*	*chr17*
cg12323063	ENSG00000267653	*Unknown*	*chr17*
cg24449629	ENSG00000243680	*Unknown*	*chr19*
cg09923107	ENSG00000264395	*MIR3193*	*chr20*
cg00494337	ENSG00000264395	*MIR3193*	*chr20*
cg17499729	ENSG00000214889	*RPS9P1*	*chr21*
cg20341238	ENSG00000188660	*LINC00319*	*chr21*
cg23454038 *	ENSG00000247993	FOXD1	*chr5*
cg08957069 *	ENSG00000221191	AL662890.1	*chr6*
cg04559779 *	ENSG00000207816	MIR124−2	chr8
cg20771240 * cg06769296*	ENSG00000254648	RP11-713P17.4	chr11
cg22771759 * cg17537073 *	ENSG00000234627	NUS1P3	*chr13*

Appendix E3: A list of probes, which are either hypermethylated (shaded grey) or hypomethylated in BBM compared to primary breast tumours and normal breast tissues. * Indicates probes located in regions other than the gene promoter.

References

Acevedo, V.D., Gangula, R.D., Freeman, K.W., Li, R.*, et al.* 2007. Inducible FGFR-1 activation leads to irreversible prostate adenocarcinoma and an epithelial-to-mesenchymal transition. *Cancer Cell,* 12**,** 559-71.

Adam, A.P., George, A., Schewe, D., Bragado, P.*, et al.* 2009. Computational identification of a p38SAPK-regulated transcription factor network required for tumor cell quiescence. *Cancer Res,* 69**,** 5664-72.

Adams, H., Fritzsche, F.R., Dirnhofer, S., Kristiansen, G. & Tzankov, A. 2010. Class I histone deacetylases 1, 2 and 3 are highly expressed in classical Hodgkin's lymphoma *Expert Opin Ther Targets.,* 14**,** 577-84.

Aguirre-Ghiso, J.A. 2007. Models, mechanisms and clinical evidence for cancer dormancy. *Nat Rev Cancer,* 7**,** 834-46.

Aguirre-Ghiso, J.A., Ossowski, L. & Rosenbaum, S.K. 2004. Green Fluorescent Protein Tagging of Extracellular Signal-Regulated Kinase and p38 Pathways Reveals Novel Dynamics of Pathway Activation during Primary and Metastatic Growth. *Cancer Research,* 64**,** 7336-45.

Ai, L., Tao, Q., Zhong, S., Fields, C.R.*, et al.* 2006. Inactivation of Wnt inhibitory factor-1 (WIF1) expression by epigenetic silencing is a common event in breast cancer. *Carcinogenesis,* 27**,** 1341-8.

Akerblom, M., Sachdeva, R., Barde, I., Verp, S.*, et al.* 2012. MicroRNA-124 is a subventricular zone neuronal fate determinant. *J Neurosci,* 32**,** 8879-89.

Akira, S. 2000. Roles of STAT3 defined by tissue-specific gene targeting. *Oncogene,* 19**,** 2607-11.

Akiyama, Y., Watkins, N., Suzuki, H., Jair, K.W.*, et al.* 2003. GATA-4 and GATA-5 Transcription Factor Genes and Potential Downstream Antitumor Target Genes Are Epigenetically Silenced in Colorectal and Gastric Cancer. *Molecular and Cellular Biology,* 23**,** 8429-39.

Albain, K.S., Allred, D.C. & Clark, G.M. 1994. Breast cancer outcome and predictors of outcome: are there age differentials? *J Natl Cancer Inst.***,** 35-42.

Alias, C., Rocchi, L., Ribatti, D., Caraffi, S.*, et al.* 2015. MMPs and angiogenesis affect the metastatic potential of a human vulvar leiomyosarcoma cell line. *J Cell Mol Med.,* Ahead of print.

American Cancer Society 2015. Global Cancer Facts and Figures 3rd Edition. *American Cancer Society*.

Antequera, F. & Bird, A. 1993. Number of CpG islands and genes in human and mouse. *Proc Natl Acad Sci U S A,* 90**,** 11995-9.

Arencibia, J.M., Martín, S., Pérez-Rodríguez, F.J. & Bonnin, A. 2009. Gene expression profiling reveals overexpression of TSPAN13 in prostate cancer. *International journal of Oncology,* 34**,** 457-63.

Banerji, S., Cibulskis, K., Rangel-Escareno, C., Brown, K.K.*, et al.* 2012. Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature,* 486**,** 405-9.

Barekati, Z., Radpour, R., Lu, Q., Bitzer, J.*, et al.* 2012. Methylation signature of lymph node metastases in breast cancer patients. *BMC Cancer,* 12**,** 244.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y.*, et al.* 2007. High-resolution profiling of histone methylations in the human genome. *Cell,* 129**,** 823-37.

Bauer, M., Dieterich, W., Tobias, E. & Detlef, S. 1997. Complete primary structure of human collagen type XIV/Undulin. *Biochimica et Biophysica Acta,* 3**,** 183-8.
Beltran, M., Puig, I., Pena, C., Garcia, J.M.*, et al.* 2008. A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. *Genes Dev,* 22**,** 756-69.
Benevolenskaya, E.V. 2007. Histone H3K4 demethylases are essential in development and differentiation. *Biochem Cell Biol.,* 85**,** 435-43.
Berger, S.L. 2007. The complex language of chromatin regulation during transcription. *Nature,* 447**,** 407-12.
Bernatchez, P.N., Allen, B.G., Gélinas, D.S., Guillemette, G. & Sirois, M.G. 2001. Regulation of VEGF-induced endothelial cell PAF synthesis: role of p42/44 MAPK, p38 MAPK and PI3K pathways. *Br J Pharmacol.,* 134**,** 1253-62.
Berois, N., Gattolliat, C.H., Barrios, E., Capandeguy, L.*, et al.* 2013. GALNT9 gene expression is a prognostic marker in neuroblastoma patients. *Clin Chem,* 59**,** 225-33.
Bibikova, M., Barnes, B., Tsan, C., Ho, V.*, et al.* 2011. High density DNA methylation array with single CpG site resolution. *Genomics,* 98**,** 288-95.
Bibikova, M., Saedinia-Melnyk, S., Zhou, L., Shen, R. & Gunderson, K. 2009. Functional normalization of 450k methylationarray data improves replication in large cancer studies. *Epigenomics,* 1**,** 177-200.
Bilguvar, K., Tyagi, N.K., Ozkara, C., Tuysuz, B.*, et al.* 2012. Recessive loss of function of the neuronal ubiquitin hydrolase UCHL1 leads to early-onset progressive neurodegeneration. *PNAS,* 110**,** 3489-94.
Bird, A. 2002. DNA methylation patterns and epigenetic memory. *Genes Dev. ,* 16**,** 6-21.
Bonazzi, V.F., Nancarrow, D.J., Stark, M.S., Moser, R.J.*, et al.* 2011. Cross-platform array screening identifies COL1A2, THBS1, TNFRSF10D and UCHL1 as genes frequently silenced by methylation in melanoma. *PLoS One,* 6**,** e26121.
Bos, P.D., Zhang, X.H., Nadal, C., Shu, W.*, et al.* 2009. Genes that mediate breast cancer metastasis to the brain. *Nature,* 459**,** 1005-9.
Bragulla, H.H. & Homberger, D.G. 2009. Structure and functions of keratin proteins in simple, stratified, keratinized and cornified epithelia. *J Anat,* 214**,** 516-59.
Brennan, C.W., Verhaak, R.G., Mckenna, A., Campos, B.*, et al.* 2013. The somatic genomic landscape of glioblastoma. *Cell,* 155**,** 462-77.
Brockhausen, I. 2006. Mucin-type O-glycans in human colon and breast cancer: glycodynamics and functions. *EMBO Rep,* 7**,** 599-604.
Burger, J.A. & Kipps, T.J. 2006. CXCR4: a key receptor in the crosstalk between tumor cells and their microenvironment. *Blood,* 107**,** 1761-7.
Burstein, H.J. & Griggs, J.J. 2012. Deep time: the long and the short of adjuvant endocrine therapy for breast cancer. *J Clin Oncol,* 30**,** 684-6.
Cancer Genome Atlas, N. 2012. Comprehensive molecular portraits of human breast tumours. *Nature,* 490**,** 61-70.
Cancer Genome Atlas Research, N. 2012. Comprehensive genomic characterization of squamous cell lung cancers. *Nature,* 489**,** 519-25.
Cancer Genome Atlas Research, N. 2013. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature,* 499**,** 43-9.
Cancer Genome Atlas Research, N. 2014. Comprehensive molecular profiling of lung adenocarcinoma. *Nature,* 511**,** 543-50.
Cancer Research Uk 2015. The cancer statistics. http://www.cancerresearchuk.org/cancer-info/cancerstats/types/skin/incidence,.

Cardoso, F., Harbeck, N., Fallowfield, L., Kyriakides, S.*, et al.* 2012. Locally recurrent or metastatic breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol,* 23 Suppl 7**,** vii11-9.

Carey, L.A., Dees, E.C., Sawyer, L., Gatti, L.*, et al.* 2007. The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. *Clin Cancer Res,* 13**,** 2329-34.

Carreres, M.I., Escalante, A., Murillo, B., Chauvin, G.*, et al.* 2011. Transcription factor Foxd1 is required for the specification of the temporal retina in mammals. *J Neurosci,* 31**,** 5673-81.

Carrio, M., Arderiu, G., Myers, C. & Boudreau, N.J. 2005. Homeobox D10 induces phenotypic reversion of breast tumor cells in a three-dimensional culture model. *Cancer Res,* 65**,** 7177-85.

Carter, C.L. & Allen, C., Henson, D.E. 1989. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. *Cancer,* 63**,** 181-7.

Cates, J.M., Dupont, W.D., Barnes, J.W., Edmunds, H.S.*, et al.* 2008. Markers of epithelial-mesenchymal transition and epithelial differentiation in sarcomatoid carcinoma: utility in the differential diagnosis with sarcoma. *Appl Immunohistochem Mol Morphol,* 16**,** 251-62.

Chai, X., Forster, E., Zhao, S., Bock, H.H. & Frotscher, M. 2009. Reelin stabilizes the actin cytoskeleton of neuronal processes by inducing n-cofilin phosphorylation at serine3. *J Neurosci,* 29**,** 288-99.

Chambers, A.F., Groom, A.C. & Macdonald, I.C. 2002. Dissemination and growth of cancer cells in metastatic sites. *Nat Rev Cancer,* 2**,** 563-72.

Chau, B.N., Diaz, R.L., Saunders, M.A., Cheng, C.*, et al.* 2009. Identification of SULF2 as a Novel Transcriptional Target of p53 by Use of Integrated Genomic Analyses. *Cancer Research,* 69**,** 1368-74.

Chaudhry, S.S., Gazzard, J., Baldock, C., Dixon, J.*, et al.* 2001. Mutation of the gene encoding fibrillin-2 results in syndactyly in mice. *Human Molecular Genetics,* 10**,** 835-43.

Chen, J., Wang, M., Guo, M., Xie, Y. & Cong, Y.S. 2013. miR-127 regulates cell proliferation and senescence by targeting BCL6. *PLoS One,* 8**,** e80266.

Chen, Y., Xu, B., Arderiu, G., Hashimoto, T.*, et al.* 2004. Retroviral delivery of homeobox D3 gene induces cerebral angiogenesis in mice. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism,* 24**,** 1280-7.

Chua, Y.J., Sargent , D., Cunningham, D. , 2005. Definitiion of disease-free survival: is this my truth-show me yours. *Annals of Oncology***,** 1719-21.

Chun, M.G. & Hanahan, D. 2010. Genetic deletion of the desmosomal component desmoplakin promotes tumor microinvasion in a mouse model of pancreatic neuroendocrine carcinogenesis. *PLoS Genet,* 6**,** e1001120.

Cianfrocca, M. & Goldstein, L., J. 2004. Prognostic and Predictive Factors in Early-Stage Breast Cancer. *The Oncologist* 9**,** 606-16.

Cihák, A. 1974. Biological effects of 5-azacytidine in eukaryotes. *Oncology* 30**,** 405-22.

Cock-Rada, A. & Weitzman, J.B. 2013. The methylation landscape of tumour metastasis. *Biol Cell,* 105**,** 73-90.

Conaco, C., Otto, S., Han, J.J. & Mandel, G. 2006. Reciprocal actions of REST and a microRNA promote neuronal identity. *Proc Natl Acad Sci U S A,* 103**,** 2422-7.

Coomber, B.L., Yu, J.L., Fathers, K.E., Plumb, C. & Rak, J.W. 2003 Angiogenesis and the role of epigenetics in metastasis. *Clinical & Experimental Metastasis,* 20**,** 215-27.

Crawford, N.P., Ziogas, A., Peel, D.J., Hess, J.*, et al.* 2006. Germline polymorphisms in SIPA1 are associated with metastasis and other indicators of poor prognosis in breast cancer. *Breast Cancer Res,* 8**,** R16.

Curtis, C., Shah, S.P., Chin, S.F., Turashvili, G.*, et al.* 2012. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature,* 486**,** 346-52.

D'atri, F., Nadalutti, F. & Citi, S. 2002. Evidence for a functional interaction between cingulin and ZO-1 in cultured cells. *J Biol Chem,* 277**,** 27757-64.

Dai, C., Tang, Y., Jung, S.Y., Qin, J.*, et al.* 2011. Differential effects on p53-mediated cell cycle arrest vs. apoptosis by p90. *Proc Natl Acad Sci U S A,* 108**,** 18937-42.

Daly, M.B., Clark, G.M. & Mcguire, W.L. 1985. Breast cancer prognosis in a mixed Caucasian-Hispanic population. *J Natl Cancer Inst.,* 74**,** 753-7.

Dammann, R., Li, C., Yoon, J.H., Chin, P.L.*, et al.* 2000. Epigenetic inactivation of a RAS association domain family protein from the lung tumour suppressor locus 3p21.3. *Nature genetics,* 25**,** 315-9.

Dammann, R., Yang, G. & Pfeifer, G.P. 2001. Hypermethylation of the cpG island of Ras association domain family 1A (RASSF1A), a putative tumor suppressor gene from the 3p21.3 locus, occurs in a large percentage of human breast cancers. *Cancer Res,* 61**,** 3105-9.

Davalos, V., Moutinho, C., Villanueva, A., Boque, R.*, et al.* 2012. Dynamic epigenetic regulation of the microRNA-200 family mediates epithelial and mesenchymal transitions in human tumorigenesis. *Oncogene,* 31**,** 2062-74.

Davies, C., Pan, H., Godwin, J., Gray, R.*, et al.* 2013. Adjuvant Tamoxifen: Longer Against Shorter (ATLAS) Collaborative Group. Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. *Lancet,* 381**,** 805-16.

Davies, E.L., Gee, J.M., Cochrane, R.A., Jiang, W.G.*, et al.* 1999. The Immunohistochemical Expression of Desmoplakin and its Role In Vivo in the Progression and Metastasis of Breast Cancer. *European Journal of Cancer,* 35**,** 902-7.

De Roos, M.A., Van Der Vegt, B., Peterse, J.L., Patriarca, C.*, et al.* 2007. The expression pattern of MUC1 (EMA) is related to tumour characteristics and clinical outcome in 'pure' ductal carcinoma in situ of the breast. *Histopathology,* 51**,** 227-38.

Deaton, A.M., Webb, S., Kerr, A.R., Illingworth, R.S.*, et al.* 2011. Cell type-specific DNA methylation at intragenic CpG islands in the immune system. *Genome Res,* 21**,** 1074-86.

Dedeurwaerder, S., Defrance, M., Calonne, E., Denis, H.*, et al.* 2011. Evaluation of the Infinium Methylation 450K technology. *Epigenomics,* 3**,** 771-84.

Degli-Esposti, M.A., Smolak, P.J., Walczak, H., Waugh, J.*, et al.* 1997. Cloning and characterization of TRAIL-R3, a novel member of the emerging TRAIL receptor family. *J Exp Med,* 186**,** 1165-70.

Dehaan, C., Habibi-Nazhad, B., Yan, E., Salloum, N.*, et al.* 2004. Mutation in mitochondrial complex I ND6 subunit is associated with defective response to hypoxia in human glioma cells. *Mol Cancer,* 3**,** 19.

Demicheli, R. 2001. Tumour dormancy: findings and hypotheses from clinical research on breast cancer. *Semin Cancer Biol,* 11**,** 297-306.

Deo, M., Yu, J.Y., Chung, K.H., Tippens, M. & Turner, D.L. 2006. Detection of mammalian microRNA expression by in situ hybridization with RNA oligonucleotides. *Dev Dyn,* 235**,** 2538-48.
Devaney, J.M., Wang, S., Funda, S., Long, J.*, et al.* 2013. Identification of novel DNA-methylated genes that correlate with human prostate cancer and high-grade prostatic intraepithelial neoplasia. *Prostate Cancer Prostatic Dis,* 16**,** 292-300.
Dhanoa, B.S., Cogliati, T., Satish, A.G., Bruford, E.A. & Friedman, J.S. 2013. Update on the Kelch-like (KLHL) gene family. *Human Genomics,* 7.
Dimmeler, S., Dernbach, E. & Zeiher, A.M. 2000. Phosphorylation of the endothelial nitric oxide synthase at ser-1177 is required for VEGF-induced endothelial cell migration. *FEBS Lett.,* 477**,** 258-62.
Disanza, A., Steffen, A., Hertzog, M., Frittoli, E.*, et al.* 2005. Actin polymerization machinery: the finish line of signaling networks, the starting point of cellular movement. *Cell Mol Life Sci.,* 62**,** 955-70.
Dong, X., Li, Y., Chang, P., Tang, H.*, et al.* 2011. Glucose Metabolism Gene Variants Modulate the Risk of Pancreatic Cancer. *Cancer Prevention Research,* 4**,** 758-66.
Du, R., Lu, K.V., Petritsch, C., Liu, P.*, et al.* 2008. HIF1alpha induces the recruitment of bone marrow-derived vascular modulatory cells to regulate tumor angiogenesis and invasion. *Cancer Cell,* 13**,** 206-20.
Dunwell, T.L., Hesson, L.B., Pavlova, T., Zabarovska, V.*, et al.* 2009b. Epigenetic analysis of childhood acute lymphoblastic leukemia. *Epigenetics : official journal of the DNA Methylation Society,* 4**,** 185-93.
Dwivedi, R.S., Herman, J.G., Mccaffrey, T.A. & Raj, D.S. 2011. Beyond genetics: epigenetic code in chronic kidney disease. *Kidney Int,* 79**,** 23-32.
Eden, A., Gaudet, F., Waghmare, A. & Jaenisch, R. 2003. Chromosomal instability and tumors promoted by DNA hypomethylation. *Science,* 300**,** 455.
Eichler, A.F., Chung, E., Kodack, D.P., Loeffler, J.S.*, et al.* 2011. The biology of brain metastases-translation to new therapies. *Nat Rev Clin Oncol,* 8**,** 344-56.
Ellis, L.M. & Fidler, I.J. 1996. Angiogenesis and metastasis. *European journal of cancer,* 32A**,** 2451-60.
Ernstsson, S., Pierrou, S., Hulander, A., Cederberg, A.*, et al.* 1996. Characterization of the Human Forkhead Gene FREAC-4. *The Journal of Biological Chemisry* 271**,** 21094-9.
Escudero-Esparza, A., Jiang, W.G. & Martin, T.A. 2012. Claudin-5 is involved in breast cancer cell motility through the N-WASP and ROCK signalling pathways. *J Exp Clin Cancer Res,* 31**,** 43.
Esteller, M., Sanchez-Cespedes, M., Rosell, R., Sidransky, D.*, et al.* 1999. Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients. *Cancer Res.,* 59 67-70.
Fang, R., Xiao, T., Fang, Z., Sun, Y.*, et al.* 2012. MicroRNA-143 (miR-143) regulates cancer glycolysis via targeting hexokinase 2 gene. *J Biol Chem,* 287**,** 23227-35.
Fazakas, C., Wilhelm, I., Nagyoszi, P., Farkas, A.E.*, et al.* 2011. Transmigration of melanoma cells through the blood-brain barrier: role of endothelial tight junctions and melanoma-released serine proteases. *PLoS One,* 6**,** e20758.
Feld, R., Rubinstein, L.V. & Weisenberger, T.H. 1984. Sites of recurrence in resected stage I non-small-cell lung cancer: a guide for future studies. *J Clin Oncol.,* 2**,** 1352-8.
Fernández-Periáñez, R., Molina-Privado, I., Rojo, F., Guijarro-Muñoz, I.*, et al.* 2013. Basement membrane-rich organoids with functional human blood vessels are permissive niches for human breast cancer metastasis. *PLoS One,* 8**,** e72957.

Ferreira, M.A., O'donovan, M.C., Meng, Y.A., Jones, I.R.*, et al.* 2008. Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nat Genet,* 40**,** 1056-8.
Feuerborn, A., Mathow, D., Srivastava, P.K., Gretz, N. & Grone, H.J. 2014. Basonuclin-1 modulates epithelial plasticity and TGF-beta1-induced loss of epithelial cell integrity. *Oncogene,* 34**,** 1185-95.
Fidler, I.J. 2003. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nat Rev Cancer,* 3**,** 453-8.
Fidler, I.J., Balasubramanian, K., Lin, Q., Kim, S.W. & Kim, S.J. 2010. The brain microenvironment and cancer metastasis. *Mol Cells,* 30**,** 93-8.
Fidler, I.J. & Poste, G. 2008. The "seed and soil" hypothesis revisited. *The Lancet Oncology,* 9**,** 808.
Fortin, J.P., Labbe, A., Lamire, M., Zanke, B.W.*, et al.* 2014. Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biology,* 15**,** 503-20.
Fritzsche, F.R., Oelrich, B., Johannsen, M., Kristiansen, I.*, et al.* 2008. Claudin-1 protein expression is a prognostic marker of patient survival in renal cell carcinomas. *Clin Cancer Res,* 14**,** 7035-42.
Fujii, H., Biel, M.A., Zhou, W., Weitzman, S.A.*, et al.* 1998. Methylation of the HIC-1 candidate tumor suppressor gene in human breast cancer. *Oncogene,* 16**,** 2159-64.
Furuta, M., Kozaki, K., Tanaka, S., Arii, S.*, et al.* 2010. miR-124 and miR-203 are epigenetically silenced tumor-suppressive microRNAs in hepatocellular carcinoma. *Carcinogenesis,* 31**,** 766-76.
Gao, Y., Chen, G., Zeng, Y., Zeng, J.*, et al.* 2015. Invasion and metastasis-related long noncoding RNA expression profiles in hepatocellular carcinoma. *Tumour Biol***,** Ahead of print.
Gardiner-Garden, M. & Frommer, M. 1987. CpG islands in vertebrate genomes. *J Mol Biol,* 196**,** 261-82.
Gavrilovic, I.T. & Posner, J.B. 2005. Brain metastases: epidemiology and pathophysiology. *Journal of neuro-oncology,* 75**,** 5-14.
Gery, S. & Koeffler, H.P. 2003. Repression of the TMEFF2 Promoter by c-Myc. *Journal of Molecular Biology,* 328**,** 977-83.
Ghosh, S.K., Pantazopoulos, P., Medarova, Z. & Moore, A. 2013. Expression of underglycosylated MUC1 antigen in cancerous and adjacent normal breast tissues. *Clin Breast Cancer,* 13**,** 109-18.
Gimenez-Roqueplo, A.P., Favier, J., Rustin, P., Mourad, J.J.*, et al.* 2001. The R22X mutation of the SDHD gene in hereditary paraganglioma abolishes the enzymatic activity of complex II in the mitochondrial respiratory chain and activates the hypoxia pathway. *Am J Hum Genet,* 69**,** 1186-97.
Giuliano, M., Giordano, A., Jackson, S., Hess, K.R.*, et al.* 2011. Circulating tumor cells as prognostic and predictive markers in metastatic breast cancer patients receiving first-line systemic treatment. *Breast Cancer Res,* 13**,** R67.
Gohji, K., Nakajima, M., Boyd, D., Dinney, C.P.*, et al.* 1997. Organ-Site Dependence for the Production of Urokinase-Type Plasminogen Activator and Metastasis by Human Renal Cell Carcinoma Cells. *Am J Pathol,* 151**,** 1655-61.
Gori, S., Rimondini, S., De Angelis, V., Colozza, M.*, et al.* 2007. Central nervous system metastases in HER-2 positive metastatic breast cancer patients treated with trastuzumab: incidence, survival, and risk factors. *Oncologist,* 12**,** 766-73.

Grady, W.M., Willis, J., Guilford, P.J., Dunbier, A.K., *et al.* 2000. Methylation of the CDH1 promoter as the second genetic hit in hereditary diffuse gastric cancer. *Nature Genetics,* 26**,** 16-7.

Greger, V., Passarge, E., Höpping, W., Messmer, E. & Horsthemke, B. 1989. Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. *Hum. Genet.,* 83**,** 155-8.

Griffith, L.G. & Swartz, M.A. 2006. Capturing complex 3D tissue physiology in vitro. *Nat Rev Mol Cell Biol,* 7**,** 211-24.

Grunstein, M., Hecht, A., Fisher-Adams, G., Wan, J., *et al.* 1995. The regulation of euchromatin and heterochromatin by histones in yeast. *J. Cell Sci Suppl.***,** 29-36.

Guo, L., Fan, D., Zhang, F., Price, J.E., *et al.* 2011. Selection of brain metastasis-initiating breast cancer cells determined by growth on hard agar. *Am J Pathol,* 178**,** 2357-66.

Gupta, G.P. & Massague, J. 2006. Cancer metastasis: building a framework. *Cell,* 127**,** 679-95.

Gvozdenovic, A., Arlt, M.J., Campanile, C., Brennecke, P., *et al.* 2013. Silencing of CD44 gene expression in human 143-B osteosarcoma cells promotes metastasis of intratibial tumors in SCID mice. *PLoS One,* 8**,** e60329.

Haffty, B.G., Yang, Q., Reiss, M., Kearney, T., *et al.* 2006. Locoregional relapse and distant metastasis in conservatively managed triple negative early-stage breast cancer. *J Clin Oncol,* 24**,** 5652-7.

Han, W., Wang, T., Reilly, A.A., Keller, S.M. & Spivack, S.D. 2009. Gene promoter methylation assayed in exhaled breath, with differences in smokers and lung cancer patients. *Respir Res,* 10**,** 86.

Han, Z.B., Yang, Z., Chi, Y., Zhang, L., *et al.* 2013. MicroRNA-124 suppresses breast cancer cell growth and motility by targeting CD151. *Cell Physiol Biochem,* 31**,** 823-32.

Hanahan, D. & Weinberg, R.A. 2000. The hallmarks of cancer. *Cell,* 100**,** 57-70.

Hanisch, F.G. 2001. O-glycosylation of the mucin type. *Biol Chem,* 382**,** 143-9.

Hansemann, D. 1890. Ueber asymmetrische Zelltheilung in epithel Krebsen und deren biologische Bedeutung. *Virchows Arch. Path. Anat,* 119.

Hanson, D., Murray, P.G., O'sullivan, J., Urquhart, J., *et al.* 2011. Exome sequencing identifies CCDC8 mutations in 3-M syndrome, suggesting that CCDC8 contributes in a pathway with CUL7 and OBSL1 to control human growth. *American journal of human genetics,* 89**,** 148-53.

Hanssens, P., Karlsson, B., Yeo, T.T., Chou, N. & Beute, G. 2011. Detection of brain micrometastases by high-resolution stereotactic magnetic resonance imaging and its impact on the timing of and risk for distant recurrences. *J Neurosurg.,* 115**,** 499-504.

Harrison, K.D., Park, E.J., Gao, N., Kuo, A., *et al.* 2011. Nogo-B receptor is necessary for cellular dolichol biosynthesis and protein N-glycosylation. *EMBO J,* 30**,** 2490-500.

Hayami, S., Kelly J.D., Cho, H.S., Yoshimatsu, M., *et al.* 2011. Overexpression of LSD1 contributes to human carcinogenesis through chromatin regulation in various cancers. *Int J Cancer,* 128**,** 574-86.

Hayashi, A., Horiuchi, A., Kikuchi, N., Hayashi, T., *et al.* 2010. Type-specific roles of histone deacetylase (HDAC) overexpression in ovarian carcinoma: HDAC1 enhances cell proliferation and HDAC3 stimulates cell migration with downregulation of E-cadherin. Int J Cancer. *Int J Cancer,* 127**,** 1332-46.

Hayes, D.F., Mesa-Tejada, R., Papsidero, L.D., Croghan, G.A., *et al.* 1991. Prediction of prognosis in primary breast cancer by detection of a high molecular weight mucin-like antigen using monoclonal antibodies DF3, F36/22, and CU18: a Cancer and Leukemia Group B study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* 9**,** 1113-23.

Herceg, Z. & Vaissiere, T. 2011. Epigenetic mechanisms and cancer: an interface between the environment and the genome. *Epigenetics,* 6**,** 804-19.

Herman, J.G., Latif, F., Weng, Y., Lerman, M.I., *et al.* 1994. Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. *Proc Natl Acad Sci,* 91**,** 9700-4.

Herman, J.G., Merlo, A., Mao, L., Lapidus, R.G., *et al.* 1995. Inactivation of the CDKN2/p16/MTS1 gene is frequently associated with aberrant DNA methylation in all common human cancers. *Cancer Res,* 55**,** 4525-30.

Hermeking, H. 2010. The miR-34 family in cancer and apoptosis. *Cell Death Differ,* 17**,** 193-9.

Hesson, L., Bièche, I., Krex, D., Criniere, E., *et al.* 204. Frequent epigenetic inactivation of RASSF1A and BLU genes located within the critical 3p21.3 region in gliomas. *Oncogene,* 23**,** 2408-19.

Hill, V.K., Ricketts, C., Bieche, I., Vacher, S., *et al.* 2011. Genome-wide DNA methylation profiling of CpG islands in breast cancer identifies novel genes associated with tumorigenicity. *Cancer Res,* 71**,** 2988-99.

Hilsenbeck, S.G., Ravdin, P.M., De Moo, R.C.A., Chamness, G.C., *et al.* 1998. Time-dependence of hazard ratios for prognostic factors in primary breast cancer. *Breast Cancer Res Treat.,* 52**,** 227-37.

Hitoshi, N., Burdon, T., Chambers, I. & Austin, S. 1998. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *GENES & DEVELOPMENT,* 12**,** 2048-60.

Hoelzinger, D.B., Demuth, T. & Berens, M.E. 2007. Autocrine factors that sustain glioma invasion and paracrine biology in the brain microenvironment. *J Natl Cancer Inst.,* 99**,** 1583-93.

Holmgren, L., O'reilly, M.S. & Folkman, J. 1995. Dormancy of micrometastases: balanced proliferation and apoptosis in the presence of angiogenesis suppression. *Nat. Med.,* 1**,** 149-53.

Hoon, D.S., Spugnardi, M., Kuo, C., Huang, S.K., *et al.* 2004. Profiling epigenetic inactivation of tumor suppressor genes in tumors and plasma from cutaneous melanoma patients. *Oncogene,* 23**,** 4014-22.

Hoque, M.O., Feng, Q., Toure, P., Dem, A., *et al.* 2006. Detection of aberrant methylation of four genes in plasma DNA for the detection of breast cancer. *J Clin Oncol,* 24**,** 4262-9.

Hsu, H.S., Chen, T.P., Hung, C.H., Wen, C.K., *et al.* 2007. Characterization of a multiple epigenetic marker panel for lung cancer detection and risk assessment in plasma. *Cancer,* 110**,** 2019-26.

Hu, G., Kang, Y. & Wang, X.F. 2009. From breast to the brain: unraveling the puzzle of metastasis organotropism. *J Mol Cell Biol,* 1**,** 3-5.

Huang, X., Wang, S., Lee, C.K., Yang, X. & Liu, B. 2011. HDAC inhibitor SNDX-275 enhances efficacy of trastuzumab in erbB2-overexpressing breast cancer cells and exhibits potential to overcome trastuzumab resistance. *Cancer Lett,* 307**,** 72-9.

Huber, A.B., Weinmann, O., Brösamle, C., Oertle, T. & Schwab, M.E. 2002. Patterns of Nogo mRNA and Protein Expression in the Developing and Adult Rat and After CNS Lesions. *J.Neurosci.,* 22**,** 3553-67.

Huber, A.B.W., O., Brosamle, C., Oertle, T., Schwab, M.E. 2002. Expression Pattern of Nogo. *J. Neurosci.,* 22**,** 3553-67.

Huber, C., Delezoide, A.L., Guimiot, F., Baumann, C*., et al.* 2009. A large-scale mutation search reveals genetic heterogeneity in 3M syndrome. *European journal of human genetics : EJHG,* 17**,** 395-400.

Huber, C., Dias-Santagata, D., Glaser, A., O'sullivan, J*., et al.* 2005. Identification of mutations in CUL7 in 3-M syndrome. *Nature genetics,* 37**,** 1119-24.

Huber, C., Munnich, A. & Cormier-Daire, V. 2011. The 3M syndrome. *Best Pract Res Clin Endocrinol Metab,* 25**,** 143-51.

Hungerford, D.A. & Nowell, P.C. 1962. Chromosome studies in human leukemia. III. Acute granulocytic leukemia. *J Natl Cancer Inst.,* 29**,** 545-65.

Hunt, S., Jones, A.V., Hinsley, E.E., Whawell, S.A. & Lambert, D.W. 2011. MicroRNA-124 suppresses oral squamous cell carcinoma motility by targeting ITGB1. *FEBS Lett,* 585**,** 187-92.

Inic, Z., Zegarac, M., Inic, M., Markovic, I*., et al.* 2014. Difference between Luminal A and Luminal B Subtypes According to Ki-67, Tumor Size, and Progesterone Receptor Negativity Providing Prognostic Information. *Clin Med Insights Oncol,* 8**,** 107-11.

Ishikawa, K., Takenaga, K., Akimoto, M., Koshikawa, N*., et al.* 2008. ROS-Generating Mitochondrial DNA Mutations Can Regulate Tumor Cell Metastasis. *Science,* 320.

Issa, J.P., Zehnbauer, B.A., Kaufmann, S.H., Biel, M.A. & Baylin, S.B. 1997. HIC1 hypermethylation is a late event in hematopoietic neoplasms. *Cancer Res,* 57**,** 1678-81.

Iyer, S.V., Dange, P.P., Alam, H., Sawant, S.S*., et al.* 2013. Understanding the role of keratins 8 and 18 in neoplastic potential of breast cancer derived cell lines. *PLoS One,* 8**,** e53532.

Jaenisch, R. & Bird, A. 2003. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet,* 33 Suppl**,** 245-54.

Jenkinson, M.D., Haylock, B., Shenoy, A., Husband, D. & Javadpour, M. 2011. Management of cerebral metastasis: evidence-based approach for surgery, stereotactic radiosurgery and radiotherapy. *European journal of cancer,* 47**,** 649-55.

Jin, C., Wang, Y., Han, W., Zhang, Y*., et al.* 2007. CMTM8 induces caspase-dependent and -independent apoptosis through a mitochondria-mediated pathway. *J Cell Physiol,* 211**,** 112-20.

Jin, K.L., Pak, J.H., Park, J.Y., Choi, W.H*., et al.* 2008. Expression profile of histone deacetylases 1, 2 and 3 in ovarian cancer tissues *J Gynecol Oncol.,* 19**,** 185-90.

Jin, L., Wessely, O., Marcusson, E.G., Ivan, C*., et al.* 2013. Prooncogenic factors miR-23b and miR-27b are regulated by Her2/Neu, EGF, and TNF-alpha in breast cancer. *Cancer Res,* 73**,** 2884-96.

Jinushi, T., Shibayama, Y., Kinoshita, I., Oizumi, S*., et al.* 2014. Low expression levels of microRNA-124-5p correlated with poor prognosis in colorectal cancer via targeting of SMC4. *Cancer Med,* 3**,** 1544-52.

Jones, P.A. & Baylin, S.B. 2007. The epigenomics of cancer. *Cell,* 128**,** 683-92.

Jones, S., Wang, T.L., Shih, L., Mao T.L., *et al.* 2010. Frequent mutations of chromatin remodeling gene ARID1A in ovarian clear cell carcinoma. *Science,* 330**,** 228-31.
Joyce, J.A. & Pollard, J.W. 2009. Microenvironmental regulation of metastasis. *Nat Rev Cancer,* 9**,** 239-52.
Kafri, T., Ariel, M., Brandeis, M., Shemer, R., *et al.* 1992. Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes Dev,* 6**,** 705-14.
Kalluri, R. & Weinberg, R.A. 2009. The basics of epithelial-mesenchymal transition. *J Clin Invest,* 119**,** 1420-8.
Kang, Y., Siegel, P.M., Shu, W., Drobnjak, M., *et al.* 2003. A multigenic program mediating breast cancer metastasis to bone. *Cancer Cell,* 3**,** 537-49.
Kar, S., Sengupta, D., Deb, M., Shilpi, A., *et al.* 2014. Expression profiling of DNA methylation-mediated epigenetic gene-silencing factors in breast cancer. *Clin Epigenetics,* 6**,** 1-13.
Kass, S.U., Pruss, D. & Wolffe, A.P. 1997. How does DNA methylation repress transcription? *Trends Genet.,* 13**,** 444-9.
Kehlen, A., Haegele, M., Menge, K., Gans, K., *et al.* 2012. Role of glutaminyl cyclases in thyroid carcinomas. *Endocr Relat cancer* 20**,** 79-90.
Kennecke, H.F., Olivotto, I.A., Speers, C., Norris, B., *et al.* 2007. Late risk of relapse and mortality among postmenopausal women with estrogen responsive early breast cancer after 5 years of tamoxifen. *Ann Oncol,* 18**,** 45-51.
Kim, T.Y., Zhong, S., Fields, C.R., Kim, J.H. & Robertson, K.D. 2006. Epigenomic profiling reveals novel and frequent targets of aberrant DNA methylation-mediated silencing in malignant glioma. *Cancer Res,* 66**,** 7490-501.
Kiyoshi, T., Koichi, N., Wei Shi, T., Tanaka, M., *et al.* 1997. Targeted disruption of the mouse Stat3 gene leads to early embryonic lethality. *proc. Natl. Acad. Sci; Developmental biology,* 94**,** 3801-4.
Klein, A., Schwartz, H., Sagi-Assif, O., Meshel, T., *et al.* 2015. Astrocytes facilitate melanoma brain metastasis via secretion of IL-23. *J Pathol.,* 236**,** 117-27.
Klein, C.A. 2003. The systemic progression of human cancer: a focus on the individual disseminated cancer cell--the unit of selection. *Adv Cancer Res,* 89**,** 35-67.
Klein, C.A. 2009. Parallel progression of primary tumours and metastases. *Nature,* 9**,** 302-12.
Koch, C.M., Andrews, R.M., Flicek, P., Dillon, S.C., *et al.* 2007. The landscape of histone modifications across 1% of the human genome in five human cell lines. *Genome Res,* 17**,** 691-707.
Koga, M., Matsuda, M., Kawamura. T., Sogo, T., *et al.* 2014. Foxd1is a mediator and indicator of the cell reprogramming process. *Nat Commun,* 5.
Koga, Y., Pelizzola, M., Cheng, E., Krauthammer, M., *et al.* 2009. Genome-wide screen of promoter methylation identifies novel markers in melanoma. *Genome Res,* 19**,** 1462-70.
Kominsky, S.L., Argani, P., Korz, D., Evron, E., *et al.* 2003. Loss of the tight junction protein claudin-7 correlates with histological grade in both ductal carcinoma in situ and invasive ductal carcinoma of the breast. *Oncogene,* 22**,** 2021-33.
Kornberg, R.D. 1974. Chromatin structure: a repeating unit of histones and DNA. *Science,* 184**,** 868-71.
Kozomara, A. & Griffiths-Jones, S. 2014. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res,* 42**,** D68-73.

Krämer, F., White, K., Kubbies, M., Swisshelm, K. & Weber, B.H. 2000. Genomic organization of claudin-1 and its assessment in hereditary and sporadic breast cancer. *Human Genetics,* 107**,** 249-56.
Kramer, N., Walzl, A., Unger, C., Rosner, M*., et al.* 2013. In vitro cell migration and invasion assays. *Mutat Res,* 752**,** 10-24.
Kuo, T.H., Kubota, T., Watanabe, M., Furukawa, T*., et al.* 1995. Liver colonization competence governs colon cancer metastasis. *Proc Natl Acad Sci,* 92**,** 12085-9.
Kurbasic, E., Sjostrom, M., Krogh, M., Folkesson, E*., et al.* 2015. Changes in glycoprotein expression between primary breast tumour and synchronous lymph node metastases or asynchronous distant metastases. *Clin Proteomics,* 12**,** 13.
Lamalice, L., Houle, F., Jourdan, G. & Huot, J. 2004. Phosphorylation of tyrosine 1214 on VEGFR2 is required for VEGF-induced activation of Cdc42 upstream of SAPK2/p38. *Oncogene,* 23**,** 434-45.
Lancet, D., Stelzer, G., Golan, Y. & Rinon, A. 2013. Gene Trends: On Muscle, Fat, and Brain. *Genetic Engineering and Biotechnology News*.
Langer, R., Mutze, K., Becker, K., Feith, M*., et al.* 2010. Expression of class I histone deacetylases (HDAC1 and HDAC2) in oesophageal adenocarcinomas: an immunohistochemical study. *J Clin Pathol.,* 63**,** 994-8.
Leong, S.P., Zuber, M., Ferris, R.L., Kitagawa, Y*., et al.* 2011. Impact of nodal status and tumor burden in sentinel lymph nodes on the clinical outcomes of cancer patients. *J Surg Oncol.,* 103**,** 518-30.
Leroy, G., Dimaggio, P.A., Chan, E.Y., Zee, B.M*., et al.* 2013. A quantitative atlas of histone modification signatures from human cancer cells. *Epigenetics Chromatin.,* 6**,** 20.
Lewis, J.D., Meehan, R.R., Henzel, W.J., Maurerfogy, I*., et al.* 1992. Purification, Sequence, and Cellular-Localization of a Novel Chromosomal Protein That Binds to Methylated DNA. *Cell,* 69**,** 905-14.
Li, E., Beard, C. & Jaenisch, R. 1993. Role for DNA methylation in genomic imprinting. *Nature,* 366**,** 362-5.
Li, Q., Qi, B., Oka, K., Shimakage, M*., et al.* 2001. Link of a new type of apoptosis-inducing gene ASY/Nogo-B to human cancer. *Oncogene,* 20**,** 3929-36.
Li, S., Butler, P., Wang, Y., Hu, Y*., et al.* 2002. The role of the dynamics of focal adhesion kinase in the mechanotaxis of endothelial cells. *Proc Natl Acad Sci,* 99**,** 3546-51.
Li, W. & Cornell, R.A. 2007. Redundant activities of Tfap2a and Tfap2c are required for neural crest induction and development of other non-neural ectoderm derivatives in zebrafish embryos. *Dev Biol,* 304**,** 338-54.
Li, X., Wang, J., Li, W., Xu, Y*., et al.* 2012. Characterization of ppGalNAc-T18, a member of the vertebrate-specific Y subfamily of UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferases. *Glycobiology,* 22**,** 602-15.
Li, Z., Pei, X.H., Yan, J., Yan, F*., et al.* 2014. CUL9 mediates the functions of the 3M complex and ubiquitylates survivin to maintain genome integrity. *Mol Cell,* 54**,** 805-19.
Lim, E., Metzger-Filho, O. & Winer, E.P. 2012. The natural history of hormone receptor-positive breast cancer. *Oncology (Williston Park),* 26**,** 688-94, 96.
Lim, S., Janzer, A., Becker, A., Zimmer, A*., et al.* 2010. Lysine-specific demethylase 1 (LSD1) is highly expressed in ER-negative breast cancers and a biomarker predicting aggressive biology. *Carcinogenesis,* 31**,** 512-20.

Lin, C., Mcgough, R., Aswad, B., Block, J.A., & Terek, R. 2004a. Hypoxia induces HIF-1alpha and VEGF expression in chondrosarcoma cells and chondrocytes. *J Orthop Res.,* 22**,** 1175-81.
Lin, N.U., Bellon, J.R. & Winer, E.P. 2004b. CNS metastases in breast cancer. *J Clin Oncol.,* 22**,** 3608-17.
Liu, S., Ren, S., Howell, P., Fodstad, O. & Riker, A.I. 2008. Identification of novel epigenetically modified genes in human melanoma via promoter methylation gene profiling. *Pigment Cell Melanoma Res,* 21**,** 545-58.
Liu, X., Shen P., Wang, X., Sun, K. & Sun, F. 2010. Solitary adrenal metastasis from invasive ductal breast cancer: an uncommon finding. *World Journal of Surgical Oncology,* 8.
Lo Nigro, C., Wang, H., Mchugh, A., Lattanzio, L.*, et al.* 2013. Methylated tissue factor pathway inhibitor 2 (TFPI2) DNA in serum is a biomarker of metastatic melanoma. *J Invest Dermatol,* 133**,** 1278-85.
Lorger, M. & Felding-Habermann, B. 2010. Capturing changes in the brain microenvironment during initial steps of breast cancer brain metastasis. *Am J Pathol,* 176**,** 2958-71.
Lorger, M., Krueger, J.S., O'neal, M., Staflin, K. & Felding-Habermann, B. 2009. Activation of tumor cell integrin alphavbeta3 controls angiogenesis and metastatic growth in the brain. *Proc Natl Acad Sci U S A,* 106**,** 10666-71.
Lowe, J.M., Menendez, D., Bushel, P.R., Shatz, M.*, et al.* 2014. p53 and NF-κB coregulate proinflammatory gene responses in human macrophages. *Cancer Res,* 74**,** 2182-92.
Loyn, M.F. 1961. Gene action in the X-chromosome of the mouse (Mus musculus L.). *Nature***,** 372-3.
Luo, H., Zhang, H., Zhang, Z., Zhang, X.*, et al.* 2009. Down-regulated miR-9 and miR-433 in human gastric carcinoma. *J Exp Clin Cancer Res,* 28**,** 82.
Ma, J., Zeng, F., Schultz, R.M. & Tseng, H. 2006. Basonuclin: a novel mammalian maternal-effect gene. *Development,* 133**,** 2053-62.
Magin, T.M., Vijayaraj, P. & Leube, R.E. 2007. Structural and regulatory functions of keratins. *Exp Cell Res,* 313**,** 2021-32.
Majid, S., Dar, A.A., Saini, S., Arora, S.*, et al.* 2012. miR-23b represses proto-oncogene Src kinase and functions as methylation-silenced tumor suppressor with diagnostic and prognostic significance in prostate cancer. *Cancer Res,* 72**,** 6435-46.
Makeyev, E.V., Zhang, J., Carrasco, M.A. & Maniatis, T. 2007. The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. *Mol Cell,* 27**,** 435-48.
Maksimova, N., Hara, K., Miyashia, A., Nikolaeva, I.*, et al.* 2007. Clinical, molecular and histopathological features of short stature syndrome with novel CUL7 mutation in Yakuts: new population isolate in Asia. *J Med Genet,* 44**,** 772-8.
Mani, S.A., Guo, W., Liao, M.J., Eaton, E.N.*, et al.* 2008. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell,* 133**,** 704-15.
Marchese, F.P. & Huarte, M. 2014. Long non-coding RNAs and chromatin modifiers: their place in the epigenetic code. *Epigenetics,* 9**,** 21-6.
Marchetti, D., Li, J. & Shen, R. 2000. Astrocytes contribute to the brain-metastatic specificity of melanoma cells by producing heparanase. *Cancer Res,* 60**,** 4767-70.

Margetts, C.D., Morris, M., Astuti, D., Gentle, D.C., *et al.* 2008. Evaluation of a functional epigenetic approach to identify promoter region methylation in phaeochromocytoma and neuroblastoma. *Endocr Relat Cancer,* 15**,** 777-86.
Marini, A., Mirmohammadsadegh, A., Nambiar, S., Gustrau, A., *et al.* 2006. Epigenetic inactivation of tumor suppressor genes in serum of patients with cutaneous melanoma. *J Invest Dermatol,* 126**,** 422-31.
Marsden, C.G., Wright, M.J., Carrier, L., Moroz, K. & Rowan, B.G. 2012. Disseminated breast cancer cells acquire a highly malignant and aggressive metastatic phenotype during metastatic latency in the bone. *PLoS One,* 7**,** e47587.
Martin, T.A., Mansel, R.E. & Jiang, W.G. 2010. Loss of occludin leads to the progression of human breast cancer. *International Journal of Molecular Medicine,* 26**,** 723-34.
Mcdonald, O.G., Wu, H., Timp, W., Doi, A. & Feinberg, A.P. 2011. Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition. *Nat Struct Mol Biol,* 18**,** 867-74.
Mcronald, F.E., Morris, M.R., Gentle, D., Winchester, L., *et al.* 2009. CpG methylation profiling in VHL related and VHL unrelated renal cell carcinoma. *Mol Cancer,* 8**,** 31.
Melki, J.R., Vincent, P.C. & Clark, S.J. 1999. Cancer-specific region of hypermethylation identified within the HIC1 putative tumour suppressor gene in acute myeloid leukaemia. *Leukemia,* 13**,** 877-83.
Mendoza, M. & Khanna, C. 2009. Revisiting the seed and soil in cancer metastasis. *Int J Biochem Cell Biol,* 41**,** 1452-62.
Metzger, F.O., Ignatiadis, M. & Sotiriou, C. 2011. Genomic Grade Index: An important tool for assessing breast cancer tumor grade and prognosis. *Crit Rev Oncol Hematol,* 77**,** 20-9.
Miller, R. & Siegmund, D. 1982. Maximally selected chi square statistics. *Biometrics***,** 1011-6.
Miloloza, A., Rosner, M., Nellist, M., Halley, D., *et al.* 2000. The TSC1 gene product hamartin, negatively regulates cell proliferation. *Human Molecular Genetics,* 9**,** 1721-7.
Minn, A.J., Gupta, G.P., Siegel, P.M., Bos, P.D., *et al.* 2005a. Genes that mediate breast cancer metastasis to lung. *Nature,* 436**,** 518-24.
Minn, A.J., Kang, Y., Serganova, I., Gupta, G.P., *et al.* 2005b. Distinct organ-specific metastatic potential of individual breast cancer cells and primary tumors. *J Clin Invest,* 115**,** 44-55.
Mittempergher, L., Saghatchian, M., Wolf, D.M., Michiels, S., *et al.* 2013. A gene signature for late distant metastasis in breast cancer identifies a potential mechanism of late recurrences. *Mol Oncol,* 7**,** 987-99.
Miyamoto, K., Asada, K., Fukutomi, T., Okochi, E., *et al.* 2003. Methylation-associated silencing of heparan sulfate D-glucosaminyl 3-O-sulfotransferase-2 (3-OST-2) in human breast, colon, lung and pancreatic cancers. *Oncogene,* 22**,** 274-80.
Mizuno, H., Kitada, K., Nakai, K. & Sarai, A. 2009. PrognoScan: a new database for meta-analysis of the prognostic value of genes. *BMC Med Genomics,* 2**,** 18.
Moelans, C.B., De Weger, R.A., Monsuur, H.N., Maes, A.H. & Van Diest, P.J. 2010. Molecular differences between ductal carcinoma in situ and adjacent invasive breast carcinoma: a multiplex ligation-dependent probe amplification study. *Anal Cell Pathol (Amst),* 33**,** 165-73.

Moelans, C.B., Verschuur-Maes, A.H. & Van Diest, P.J. 2011. Frequent promoter hypermethylation of BRCA2, CDH13, MSH6, PAX5, PAX6 and WT1 in ductal carcinoma in situ and invasive breast cancer. *J Pathol,* 225**,** 222-31.

Monk, M., Boubelik, M. & Lehnert, S. 1987. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development,* 99**,** 371-82.

Moreno, D.A., Scrideli, C.A., Cortez, M.A., De Paula, Q.R.*, et al.* 2010. Differential expression of HDAC3, HDAC7 and HDAC9 is associated with prognosis and survival in childhood acute lymphoblastic leukaemia. *Br J Haematol,* 150**,** 665-73.

Morris, M.R., Gentle, D., Abdulrahman, M., Clarke, N.*, et al.* 2008. Functional epigenomics approach to identify methylated candidate tumour suppressor genes in renal cell carcinoma. *Br J Cancer,* 98**,** 496-501.

Morris, M.R., Gentle, D., Abdulrahman, M., Maina, E.N.*, et al.* 2005. Tumor suppressor activity and epigenetic inactivation of hepatocyte growth factor activator inhibitor type 2/SPINT2 in papillary and clear cell renal cell carcinoma. *Cancer Res,* 65**,** 4598-606.

Morris, M.R., Ricketts, C., Gentle, D., Abdulrahman, M.*, et al.* 2010. Identification of candidate tumour suppressor genes frequently methylated in renal cell carcinoma. *Oncogene,* 29**,** 2104-17.

Morris, M.R., Ricketts, C.J., Gentle, D., Mcronald, F.*, et al.* 2011. Genome-wide methylation analysis identifies epigenetically inactivated candidate tumour suppressor genes in renal cell carcinoma. *Oncogene,* 30**,** 1390-401.

Morrissey, C., Martinez, A., Zatyka, M., Agathanggelou, A.*, et al.* 2001. Epigenetic inactivation of the RASSF1A 3p21.3 tumor suppressor gene in both clear cell and papillary renal cell carcinoma. *Cancer Res,* 61**,** 7277-81.

Muggerud, A.A., Ronneberg, J.A., Warnberg, F., Botling, J.*, et al.* 2010. Frequent aberrant DNA methylation of ABCB1, FOXC1, PPP2R2B and PTEN in ductal carcinoma in situ and early invasive breast cancer. *Breast cancer research : BCR,* 12**,** R3.

Muhlisch, J., Schwering, A., Grotzer, M., Vince, G.H.*, et al.* 2006. Epigenetic repression of RASSF1A but not CASP8 in supratentorial PNET (sPNET) and atypical teratoid/rhabdoid tumors (AT/RT) of childhood. *Oncogene,* 25**,** 1111-7.

Müller, A., Homey, B., Soto, H., Ge, N.*, et al.* 2001. Involvement of chemokine receptors in breast cancer metastasis. *Nature,* 410**,** 50-6.

Mulvihill, M.S., Kwon, Y.W., Lee, S., Fang, L.T.*, et al.* 2012. Gremlin is overexpressed in lung adenocarcinoma and increases cell growth and proliferation in normal lung cells. *PLoS One,* 7**,** e42264.

Mundy, G.R. 2002. Metastasis to bone: causes, consequences and therapeutic opportunities. *Nat Rev Cancer,* 2**,** 584-93.

Munster, P.N., Thurn, K.T., Thomas, S., Raha, P.*, et al.* 2011. A phase II study of the histone deacetylase inhibitor vorinostat combined with tamoxifen for the treatment of patients with hormone therapy-resistant breast cancer. *Br J Cancer,* 104**,** 1828-35.

Murata, J., Ricciardi-Castagnoli, P., Dessous, L., Mange, P.*, et al.* 1997. Microglial cells induce cytotoxic effects toward colon carcinoma cells: measurement of tumor cytotoxicity with a gamma-glutamyl transpeptidase assay. *Intl J Cancer,* 70**,** 169-74.

Murray, P.G., Hanson, D., Coulson, T., Stevens, A.*, et al.* 2013. 3-M syndrome: a growth disorder associated with IGF2 silencing. *Endocrine connections,* 2**,** 1-11.

Mutze, K., Langer, R., Becker, K., Ott, K.*, et al.* 2010. Histone deacetylase (HDAC) 1 and 2 expression and chemotherapy in gastric cancer. *Ann Surg Oncol.,* 17**,** 3336-43.

Myal, Y., Leygue, E. & Blanchard, A.A. 2010. Claudin 1 in breast tumorigenesis: revelation of a possible novel "claudin high" subset of breast cancers. *J Biomed Biotechnol,* 2010**,** 956897.

Nakagawa, S., Deli, M.A., Kawaguchi, H., Shimizudani, T.*, et al.* 2009. A new blood-brain barrier model using primary rat brain endothelial cells, pericytes and astrocytes. *Neurochem Int.,* 54**,** 253-63.

Nakajima, M., Morikawa, K., Fabra, A., Bucana, C.D. & Fidler, I.J. 1990. Influence of Organ Environment on Extracellular Matrix Degradative Activity and Metastasis of Human Colon Carcinoma Cells. *Journal of National Cancer Institute***,** 1890-8.

Neman, J., Choy, C., Kowolik, C.M., Anderson, A.*, et al.* 2013. Co-evolution of breast-to-brain metastasis and neural progenitor cells. *Clin Exp Metastasis,* 30**,** 753-68.

Neman, J., Termini, J., Wilczynski, S., Vaidehi, N.*, et al.* 2014. Human breast cancer metastases to the brain display GABAergic properties in the neural niche. *Proc Natl Acad Sci U S A,* 111**,** 984-9.

Neo, W.H., Yap, K., Lee, S.H., Looi, L.S.*, et al.* 2014. MicroRNA miR-124 controls the choice between neuronal and astrocyte differentiation by fine-tuning Ezh2 expression. *J Biol Chem,* 289**,** 20788-801.

Neve, R.M., Chin, K., Fridlyand, J., Yeh, J.*, et al.* 2006. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell,* 10**,** 515-27.

Nguyen, D.X., Bos, P.D. & Massague, J. 2009. Metastasis: from dissemination to organ-specific colonization. *Nat Rev Cancer,* 9**,** 274-84.

Nguyen, D.X. & Massague, J. 2007. Genetic determinants of cancer metastasis. *Nat Rev Genet,* 8**,** 341-52.

Nitta, T., Hata, M., Gotoh, S., Seo, Y.*, et al.* 2003. Size-selective loosening of the blood-brain barrier in claudin-5-deficient mice. *J Cell Biol,* 161**,** 653-60.

Nomoto, S., Nakao, A., Ando, N., Takeda, S.*, et al.* 1998. Clinical application of K-ras oncogene mutations in pancreatic carcinoma: detection of micrometastases. *Semin Surg Oncol.,* 15**,** 40-6.

O'leary, V.B., Ovsepian, S.V., Carrascosa, L.G., Buske, F.A.*, et al.* 2015. PARTICLE, a Triplex-Forming Long ncRNA, Regulates Locus-Specific Methylation in Response to Low-Dose Irradiation. *Cell Rep*.

Oertle, T., Merkler, D. & Schwab, M.E. 2003. Do cancer cells die because of Nogo-B? *Oncogene,* 22**,** 1390-9.

Okano, M., Bell, D.W., Haber, D.A. & Li, E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell,* 99**,** 247-57.

Olsson, A.K., Dimberg, A., Kreuger, J. & Claesson-Welsh, L. 2006. VEGF receptor signalling in control of vascular function. *Nat Rev Mol Cell Biol.,* 7**,** 359-71.

Onitilo, A.A., Engel, J.M., Greenlee, R.T. & Mukesh, B.N. 2009. Breast cancer subtypes based on ER/PR and Her2 expression: comparison of clinicopathologic features and survival. *Clin Med Res,* 7**,** 4-13.

Osanai, M., Murata, M., Chiba, H., Kojima, T. & Sawada, N. 2007. Epigenetic silencing of claudin-6 promotes anchorage-independent growth of breast carcinoma cells. *Cancer Sci,* 98**,** 1557-62.

Oshima, R.G. 2002. Apoptosis and keratin intermediate filaments. *Cell Death and Differentiation,* 9**,** 486-92.
Osler, M.E., Smith, T.K. & Bader, D.M. 2006. Bves, a member of the Popeye domain-containing gene family. *Dev Dyn,* 235**,** 586-93.
Ouyang, P. & Sugrue, S.P. 1996. Characterization of Pinin, A Novel Protein Associated with the Desmosome-Intermediate Filament Complex. *The Journal of Cell Biology,* 135**,** 1027-42.
Paget, M.R. 1889. The distribution of secondary growths in cancer of the breast. *The Lancet* 571-3.
Paik, S., Shak, S., Tang, G., Kim, C.*, et al.* 2004. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med,* 351**,** 2817-26.
Paliwal, A., Vaissiere, T., Krais, A., Cuenin, C.*, et al.* 2010. Aberrant DNA methylation links cancer susceptibility locus 15q25.1 to apoptotic regulation and lung cancer. *Cancer Res,* 70**,** 2779-88.
Pang, H., Rowan, B.G., Al-Dhaheri, M. & Faber, L.E. 2004. Epidermal growth factor suppresses induction by progestin of the adhesion protein desmoplakin in T47D breast cancer cells. *Breast Cancer Res,* 6**,** R239-45.
Pangeni, R.P., Channathodiyil, P., Huen, D.S., Eagles, L.W.*, et al.* 2015. The GALNT9, BNC1 and CCDC8 genes are frequently epigenetically dysregulated in breast tumours that metastasise to the brain. *Clinical Epigenetics,* 7**,** 1-17.
Pantel, K., Brakenhoff, R.H. & Brandt, B. 2008. Detection, clinical relevance and specific biological properties of disseminating tumour cells. *Nat Rev Cancer,* 8**,** 329-40.
Park, B.K., Zhang, H., Zeng, Q., Dai, J.*, et al.* 2007. NF-kappaB in breast cancer cells promotes osteolytic bone metastasis by inducing osteoclastogenesis via GM-CSF. *Nat Med,* 13**,** 62-9.
Park, S.Y., Kwon, H.J., Lee, H.E., Ryu, H.S.*, et al.* 2011. Promoter CpG island hypermethylation during breast cancer progression. *Virchows Arch,* 458**,** 73-84.
Patel, R., Shervington, L., Lea, R. & Shervington, A. 2008. Epigenetic silencing of telomerase and a non-alkylating agent as a novel therapeutic approach for glioma. *Brain Res,* 1188**,** 173-81.
Patra, S.K. & Bettuzzi, S. 2009. Epigenetic DNA-(cytosine-5-carbon) modifications: 5-aza-2'-deoxycytidine and DNA-demethylation. *Biochemistry (Mosc),* 74**,** 613-9.
Pellegrino, L., Stebbing, J., Braga, V.M., Frampton, A.E.*, et al.* 2013. miR-23b regulates cytoskeletal remodeling, motility and metastasis by directly targeting multiple transcripts. *Nucleic Acids Res,* 41**,** 5400-12.
Peng, C., Togayachi,A., Kwon,Y., Xie, C. Wu G , Zou,X., Sato, T.,Ito,H., Tachibana,K., Kubota,T., Noce,T., Narimatsu,H., Zhang,Y. 2010. Identification of a novel human UDP-GalNAc transferase with unique catalytic activity and expression profile. *Biochemical and Biophysical Research Communications***,** 680-6.
Perey, L., Hayes, D.F. & Kufe, D. 1992. Effects of differentiating agents on cell surface expression of the breast carcinoma-associated DF3-P epitope. *Cancer research,* 52**,** 6365-70.
Perini, G., Diolaiti, D., Porro, A. & Della Valle, G. 2005. In vivo transcriptional regulation of N-Myc target genes is controlled by E-box methylation. *Proc Natl Acad Sci U S A,* 102**,** 12117-22.
Pestalozzi, B.C. 2009. Brain metastases and subtypes of breast cancer. *Ann Oncol,* 20**,** 803-5.

Pestalozzi, B.C., Zahrieh, D., Price, K.N., Holmberg, S.B., *et al.* 2006. International Breast Cancer Study Group (IBCSG). Identifying breast cancer patients at risk for Central Nervous System (CNS) metastases in trials of the International Breast Cancer Study Group (IBCSG). *Ann Oncol.,* 17**,** 935-44.

Poliseno, L., Salmena, L., Zhang, J., Carver, B., *et al.* 2010. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature,* 465**,** 1033-8.

Polyak, K. & Weinberg, R.A. 2009. Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. *Nat Rev Cancer,* 9**,** 265-73.

Ponomarev, E.D., Veremeyko, T., Barteneva, N., Krichevsky, A.M. & Weiner, H.L. 2011. MicroRNA-124 promotes microglia quiescence and suppresses EAE by deactivating macrophages via the C/EBP-alpha-PU.1 pathway. *Nat Med,* 17**,** 64-70.

Pradhan, S., Bacolla, A., Wells, R.D. & Roberts, R.J. 1999. Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation. *J Biol Chem,* 274**,** 33002-10.

Psaila, B. & Lyden, D. 2009. The metastatic niche: adapting the foreign soil. *Nature Reviews,* 9**,** 285-93.

Pulling, L.C., Grimes, M.J., Damiani, L.A., Juri, D.E., *et al.* 2009. Dual promoter regulation of death-associated protein kinase gene leads to differentially silenced transcripts by methylation in cancer. *Carcinogenesis,* 30**,** 2023-30.

Quina, A.S., Buschbeck, M. & Di Croce, L. 2006. Chromatin structure and epigenetics. *Biochem Pharmacol.,* 72**,** 1563-9.

Raman, J., Guan, Y., Perrine, C.L., Gerken, T.A. & Tabak, L.A. 2012. UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferases: completion of the family tree. *Glycobiology,* 22**,** 768-77.

Rao-Bindal, K. & Kleinerman, E.S. 2011. Epigenetic regulation of apoptosis and cell cycle in osteosarcoma. *Sarcoma,* 2011**,** 679457.

Rappaport, N., Twik, M., Nativ, N., Stelzer, G., *et al.* 2014. MalaCards: A Comprehensive Automatically-Mined Database of Human Diseases. *Current Protocols in Bioinformatics,* 47:1.24:1.24.1–1.24.19.

Reddy, E.P., Reynolds, R.K., Santos, E. & Barbacid, M. 1982. A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma. *Oncogene,* 300**,** 149-52.

Regel, I., Eichenmuller, M., Joppien, S., Liebl, J., *et al.* 2012. IGFBP3 impedes aggressive growth of pediatric liver cancer and is epigenetically silenced in vascular invasive and metastatic tumors. *Molecular cancer,* 11**,** 9.

Regier, D.S., Higbee, J., Lund, K.M., Sakane, F., *et al.* 2005. Diacylglycerol kinase iota regulates Ras guanyl-releasing protein 3 and inhibits Rap1 signaling. *Proc Natl Acad Sci U S A,* 102**,** 7595-600.

Riethdorf, S., Fritsche, H., Muller, V., Rau, T., *et al.* 2007. Detection of circulating tumor cells in peripheral blood of patients with metastatic breast cancer: a validation study of the CellSearch system. *Clin Cancer Res,* 13**,** 920-8.

Roberts, T.C., Morris, K.V. & Weinberg, M.S. 2014. Perspectives on the mechanism of transcriptional regulation by long non-coding RNAs. *Epigenetics,* 9**,** 13-20.

Rocken, M. 2010. Early tumor dissemination, but late metastasis: insights into tumor dormancy. *J Clin Invest,* 120**,** 1800-3.

Rodenhiser, D.I. 2009. Epigenetic contributions to cancer metastasis. *Clin Exp Metastasis,* 26**,** 5-18.

Rodenhiser, D.I., Andrews, J., Kennette, W., Sadikovic, B.*, et al.* 2008. Epigenetic mapping and functional analysis in a breast cancer metastasis model using whole-genome promoter tiling microarrays. *Breast Cancer Res,* 10**,** R62.

Rodriguez-Paredes, M. & Esteller, M. 2011. Cancer epigenetics reaches mainstream oncology. *Nat Med,* 17**,** 330-9.

Rokhlin, O.W., Scheinker, V.S., Taghiyev, A.F., Bumcrot, D.*, et al.* 2008. MicroRNA-34 mediates AR-dependent p53-induced apoptosis in prostate cancer. *Cancer Biol Ther,* 7**,** 1288-96.

Roloff, T.C., Ropers, H.H. & Nuber, U.A. 2003. Comparative study of methyl-CpG-binding domain proteins. *BMC Genomics,* 4**,** 1.

Ronneberg, J.A., Fleischer, T., Solvang, H.K., Nordgard, S.H.*, et al.* 2011. Methylation profiling with a panel of cancer related genes: association with estrogen receptor, TP53 mutation status and expression subtypes in sporadic breast cancer. *Mol Oncol,* 5**,** 61-76.

Rosen, H., Krichevsky. A., Polakiewicz, R.D., Benzakine, S. & Bar-Shavit, Z. 2013b. Developmental regulation of proenkephalin gene expression in osteoblasts. *Molecular Endocrinology,* 9**,** 1621-31.

Rosenfeld, J.A., Wang, Z., Schones, D.E., Zhao, K.*, et al.* 2009. Determination of enriched histone modifications in non-genic portions of the human genome. *BMC Genomics,* 10**,** 143.

Rotman, G. & Shiloh, Y. 1998. ATM: from gene to function. *Human Molecular Genetics,* 7**,** 1555-63.

Rousseau, S., Houle, F. & Huot, J. 2000. Integrating the VEGF signals leading to actin-based motility in vascular endothelial cells. *Trends Cardiovasc Med.,* 10**,** 321-7.

Roy, S., Levi, E., Majumdar, A.P.N. & Sarkar, F.H. 2012. Expression of miR-34 is lost in colon cancer which can be re-expressed by a novel agent CDF. *Journal of Hematology and Oncology,* 5**,** 58-63.

Ryberg, M., Nielsen, D., Osterlind, K., Andersen, P.K.*, et al.* 2005. Predictors of central nervous system metastasis in patients with metastatic breast cancer. A competing risk analysis of 579 patients treated with epirubicin-based chemotherapy. *Breast Cancer Res Treat,* 91**,** 217-25.

Sabine, S., Karen, G., Qiu, M.L., Heather, G. & Marc, L. 1998. Characterization of Two Novel Protocadherins (PCDH8 and PCDH9) Localized on Human Chromosome 13 and Mouse Chromosome 14. *GENOMICS,* 53**,** 81-9.

Saez, R.A., Mcguire, W.L. & Clark, G.M. 1989. Prognostic factors in breast cancer. *Semn Surg Oncol.,* 5**,** 102-10.

Safran, M., Dalah, I., Alexander, J., Rosen, N.*, et al.* 2010. GeneCards Version 3: the human gene integrator. *Database (Oxford),* 2010**,** baq020.

Sakai, T., Toguchida, J., Ohtani, N., Yandell, D.W.*, et al.* 1991. Allele-specific hypermethylation of the retinoblastoma tumor-suppressor gene. *Am J Hum Genet,* 48**,** 880-8.

Sakuma, T., Uzawa, K., Onda, T., Shiiba, M.*, et al.* 2006. Aberrant expression of histone deacetylase 6 in oral squamous cell carcinoma. *Int J Oncol.,* 29**,** 117-24.

Sakurai, T. 2012. The role of NrCAM in neural development and disorders--beyond a simple glue in the brain. *Mol Cell Neurosci,* 49**,** 351-63.

Salhia, B., Kiefer, J., Ross, J.T., Metapally, R.*, et al.* 2014. Integrated genomic and epigenomic analysis of breast cancer brain metastasis. *PLoS One,* 9**,** e85448.

Sanna, G., Franceschelli, L., Rotmensz, N., Botteri, E.*, et al.* 2007. Brain metastases in patients with advanced breast cancer. *Anticancer Res,* 27**,** 2865-9.

Sarasqueta, A.F., Forte, F.I., Corver, W.E., De Miranda, N.F., *et al.* 2013. Integral analysis of p53 and its value as prognostic factor in sporadic colon cancer. *BMC cancer*, 277-87.
Sarrio, D., Rodriguez-Pinilla, S.M., Hardisson, D., Cano, A., *et al.* 2008. Epithelial-mesenchymal transition in breast cancer relates to the basal-like phenotype. *Cancer Res,* 68, 989-97.
Sawan, C. & Herceg, Z. 2010. Histone modifications and cancer. *Adv Genet*, 57-85.
Schaefer, M., Steringer, J.P. & Lyko, F. 2008. The Drosophila cytosine-5 methyltransferase Dnmt2 is associated with the nuclear matrix and can access DNA during mitosis. *PLoS One,* 3, e1414.
Schinke, C., Mo, Y., Yu, Y., Amiri, K., *et al.* 2010. Aberrant DNA methylation in malignant melanoma. *Melanoma Res,* 20, 253-65.
Schinzel, A.C. & Hahn, W.C. 2008. Oncogenic transformation and experimental models of human cancer. *Front Biosci.,* 13, 71-84.
Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods.,* 9, 671-5.
Schulte, J.H., Lim, S., Schramm, A., Friedrichs, N., *et al.* 2009. Lysine-specific demethylase 1 is strongly expressed in poorly differentiated neuroblastoma: implications for therapy. *Cancer Res,* 69, 2065.
Schwab, M.E. 2010. Functions of Nogo proteins and their receptors in the nervous system. *Nat Rev Neurosci,* 11, 799-811.
Senger, D.R., Ledbetter, S.R., Claffey, K.P., Papadopoulos-Sergiou, A., *et al.* 1996. Stimulation of endothelial cell migration by vascular permeability actor/vascular endothelial growth factor through cooperative mechanisms involving the alphavbeta3 integrin, osteopontin, and thrombin. *Am J Pathol.,* 149, 293-305.
Shames, D.S., Girard, L., Gao, B., Sato, M., *et al.* 2006b. A genome-wide screen for promoter methylation in lung cancer identifies novel methylation markers for multiple malignancies. *PLoS Med,* 3, e486.
Shao, C., Sun, W., Tan, M., Glazer, C.A., *et al.* 2011. Integrated, genome-wide screening for hypomethylated oncogenes in salivary gland adenoid cystic carcinoma. *Clin Cancer Res,* 17, 4320-30.
Shenker, N.S., Flower, K.J., Wilhelm-Benartzi, C.S., Dai, W., *et al.* 2015. Transcriptional implications of intragenic DNA methylation in the oestrogen receptor alpha gene in breast cancer cells and tissues. *BMC Cancer,* 15, 337.
Shi, X.B., Xue, L., Ma, A.H., Tepper, C.G., *et al.* 2013. Tumor suppressive miR-124 targets androgen receptor and inhibits proliferation of prostate cancer cells. *Oncogene,* 32, 4130-8.
Shimomura, Y., Wajid, M., Kurban, M., Sato, N. & Christiano, A.M. 2010. Mutations in the keratin 85 (KRT85/hHb5) gene underlie pure hair and nail ectodermal dysplasia. *J Invest Dermatol,* 130, 892-5.
Shinya, T., Mari, T., Morichika, K. & Tadahisa, M. 2000. Brain-speci¢c expression of a novel human UDP-GalNAc:polypeptide
N-acetylgalactosaminyltransferase (GalNAc-T9). *Biochimica et Biophysica Acta,* 1493, 264-8.
Shiozaki, A., Nakashima, S., Ichikawa, D., Fujiwara, H., *et al.* 2013. Prognostic significance of p21 expression in patients with esophageal squamous cell carcinoma. *Anticancer Res,* 10, 4329-35.
Siraj, A.K., Hussain, A.R., Al-Rasheed, M., Ahmed, M., *et al.* 2011. Demethylation of TMS1 gene sensitizes thyroid cancer cells to TRAIL-induced apoptosis. *The Journal of clinical endocrinology and metabolism,* 96, E215-24.

Slamon, D.J., Leyland-Jones, B., Shak, S., Fuchs, H., *et al.* 2001. USE OF CHEMOTHERAPY PLUS A MONOCLONAL ANTIBODY AGAINST HER2 FOR METASTATIC BREAST CANCER THAT OVEREXPRESSES HER2. *The New England Journal of Medicine,* 344**,** 783-92.

Sleeman, J.P. 2012. Metastasis: understanding is the beginning of order in chaos. *Semin Cancer Biol,* 22**,** 173.

Soffietti, R., Ruda, R. & Mutani, R. 2002. Management of brain metastases. *Journal of neurology,* 249**,** 1357-69.

Soffietti, R., Ruda, R. & Trevisan, E. 2008. Brain metastases: current management and new developments. *Current opinion in oncology,* 20**,** 676-84.

Solaroli, N., Panayiotou, C., Johansson, M. & Karlsson, A. 2009. Identification of two active functional domains of human adenylate kinase 5. *FEBS Lett,* 583**,** 2872-6.

Somiari, S.B., Somiari, R.I., Heckman, C.M., Olsen, C.H., *et al.* 2006. Circulating MMP2 and MMP9 in breast cancer -- potential role in classification of patients into low risk, high risk, benign disease and breast cancer categories. *International journal of cancer. Journal international du cancer,* 119**,** 1403-11.

Sorlie, T., Tibshirani, R., Parker, J., Hastie, T., *et al.* 2003. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A,* 100**,** 8418-23.

Spano, D., Heck, C., De Antonellis, P., Christofori, G. & Zollo, M. 2012. Molecular networks that regulate cancer metastasis. *Semin Cancer Biol,* 22**,** 234-49.

Spellmann, I., Rujescu, D., Musil, R., Giegling, I., *et al.* 2014. Pleckstrin homology domain containing 6 protein (PLEKHA6) polymorphisms are associated with psychopathology and response to treatment in schizophrenic patients. *Prog Neuropsychopharmacol Biol Psychiatry,* 51**,** 190-5.

Sperduto, P.W., Chao, S.T., Sneed, P.K., Luo, X., *et al.* 2010. Diagnosis-specific prognostic factors, indexes, and treatment outcomes for patients with newly diagnosed brain metastases: a multi-institutional analysis of 4,259 patients. *International journal of radiation oncology, biology, physics,* 77**,** 655-61.

Stark, A.M., Tongers, K., Nicolai, M.H., Maximilian, M. & Held-Feindt, J. 2005. Reduced metastasis-suppressor gene mRNA-expression in breast cancer brain metastases. *Journal of Cancer Research and Clinical Oncology,* 131**,** 191-8.

Stark, M.S., Tyagi, S., Nancarrow, D.J., Boyle, G.M., *et al.* 2010. Characterization of the Melanoma miRNAome by Deep Sequencing. *PLoS One,* 5**,** e9685.

Steger, D.J., Lefterova, M.I., Ying, L., Stonestrom, A.J., *et al.* 2008. DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells. *Mol Cell Biol,* 28**,** 2825-39.

Stephens, P.J., Tarpey, P.S., Davies, H., Van Loo, P., *et al.* 2012. The landscape of cancer genes and mutational processes in breast cancer. *Nature,* 486**,** 400-4.

Stratton, M.R., Campbell, P.J. & Futreal, P.A. 2009. The cancer genome. *Nature,* 458**,** 719-24.

Subramani, A., Alsidawi, S., Jagannathan, S., Sumita, K., *et al.* 2013. The brain microenvironment negatively regulates miRNA-768-3p to promote K-ras expression and lung cancer metastasis. *Scientific Reports,* 3**,** 1-11.

Sugarbaker, E.D. 1952. The organ selectivity of experimentally induced metastases in rats. *Cancer,* 5**,** 606-12.

Sugiyama, Y., Kusuhara, H. & Suzuki, H. 1999. Kinetic and biochemical analysis of carrier-mediated efflux of drugs through the blood-brain and blood-cerebrospinal fluid barriers: importance in the drug delivery to the brain. *J Control Release.,* 62**,** 179-86.

Symmans, W.F. 2010. Tissue banking. *Breast Cancer Research,* 12**,** L1.
Takahashi, H., Sakuta, H., Shintani, T. & Noda, M. 2009. Functional mode of FoxD1/CBF2 for the establishment of temporal retinal specificity in the developing chick retina. *Dev Biol,* 331**,** 300-10.
Talmadge, J.E. & Fidler, I.J. 2010. AACR centennial series: the biology of cancer metastasis: historical perspective. *Cancer Res,* 70**,** 5649-69.
Taube, J.H., Herschkowitz, J.I., Komurov, K., Zhou, A.Y*., et al.* 2010. Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. *Proc Natl Acad Sci,* 107**,** 15449-54.
Tcga 2012. Comprehensive molecular portraits of human breast tumours. *Nature,* 490**,** 61-70.
Termini, J., Neman, J. & Jandial, R. 2014. Role of the neural niche in brain metastatic cancer. *Cancer Res,* 74**,** 4011-5.
Teschemacher, A.G., Gourine, A.V. & Kasparov, S. 2015. A Role for Astrocytes in Sensing the Brain Microenvironment and Neuro-Metabolic Integration. *Neurochem Res.,* Ahead of print.
Thakur, A., Rahman, K.W., Wu, J., Bollig, A*., et al.* 2007. Aberrant expression of X-linked genes RbAp46, Rsk4, and Cldn2 in breast cancer. *Mol Cancer Res,* 5**,** 171-81.
Tham, Y.L., Sexton, K., Kramer, R., Hilsenbeck, S. & Elledge, R. 2006. Primary breast cancer phenotypes associated with propensity for central nervous system metastases. *Cancer,* 107**,** 696-704.
Theriot, A.J. & Mitchison, T.J. 1991. Actin microfilaments dynamics in locomotion cells *Nature,* 352.
Thiery, J.P. 2002. Epithelial-mesenchymal transitions in tumour progression. *Nat Rev Cancer,* 2**,** 442-54.
Thiery, J.P. & Sleeman, J.P. 2006. Complex networks orchestrate epithelial-mesenchymal transitions. *Nat Rev Mol Cell Biol,* 7**,** 131-42.
Thomson, S., Petti, F., Sujka-Kwok, I., Mercado, P*., et al.* 2011. A systems view of epithelial-mesenchymal transition signaling states. *Clin Exp Metastasis,* 28**,** 137-55.
Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J*., et al.* 2012. The accessible chromatin landscape of the human genome. *Nature,* 489**,** 75-82.
Toba, S., Tenno, M., Konishi, M., Mikami, T*., et al.* 2000. Brain-specifc expression of a novel human UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase (GalNAc T9). *Biochimica et Biophysica Acta***,** 264-8.
Todd, S.C., Doctor, V.S. & Levy, S. 1998a. Sequences and expression of six new members of the tetraspanin/TM4SF family. *Biochem Biophys Acta,* 1**,** 101-4.
Todd, S.C., Doctor, V.S. & Levy, S. 1998b. Sequences and expression of six new members of the tetraspanin/TM4SF family. 101-4.
Tommasi, S., Karm, D.L., Wu, X., Yen, Y. & Pfeifer, G.P. 2009. Methylation of homeobox genes is a frequent and early epigenetic event in breast cancer. *Breast Cancer Res,* 11**,** R14.
Torre, L.A., Bray, F., Siegel, R.L., Ferlay, J*., et al.* 2015. Global cancer statistics, 2012. *CA Cancer J Clin,* 65**,** 87-108.
Tosoni, A., Franceschi, E. & Brandes, A.A. 2008. Chemotherapy in breast cancer patients with brain metastases: have new chemotherapic agents changed the clinical outcome? *Crit Rev Oncol Hematol,* 68**,** 212-21.

Toyooka, S., Toyooka, K.O., Miyajima, K., Reddy, J.L., *et al.* 2003. Epigenetic Down-Regulation of Death-associated Protein Kinase in Lung Cancers. *Clinical Cancer Research,* 9**,** 3034-41.
Tsai, H.C. & Baylin, S.B. 2011. Cancer epigenetics: linking basic biology to clinical medicine. *Cell Res,* 21**,** 502-17.
Tsukita, S. & Furuse, M. 2000. The Structure and Function of Claudins, Cell Adhesion Molecules at Tight Junctions. *Ann N Y Acad Sci,* 915**,** 129-35.
Turksen, K. & Troy, T.C. 2004. Barriers built on claudins. *J Cell Sci,* 117**,** 2435-47.
Tuupanen, S., Hanninen, U.A., Kondelin, J., Von Nandelstadh, P., *et al.* 2014. Identification of 33 candidate oncogenes by screening for base-specific mutations. *Br J Cancer,* 111**,** 1657-62.
Ushio-Fukai, M., Tang, Y., Fukai, T., Dikalov, S.I., *et al.* 2002. Novel role of gp91(phox)-containing NAD(P)H oxidase in vascular endothelial growth factor-induced signaling and angiogenesis. *Circ. Res.,* 91**,** 1160-7.
Valastyan, S. & Weinberg, R.A. 2011. Tumor metastasis: molecular insights and evolving paradigms. *Cell,* 147**,** 275-92.
Van 'T Veer, L.J., Dai, H., Van De Vijver, M.J., He, Y.D., *et al.* 2002. Gene expression profiling predicts clinical outcome of breast cancer. *Nature,* 415**,** 530-6.
Van Der Vegt, B., De Roos, M.A., Peterse, J.L., Patriarca, C., *et al.* 2007. The expression pattern of MUC1 (EMA) is related to tumour characteristics and clinical outcome of invasive ductal breast carcinoma. *Histopathology,* 51**,** 322-35.
Van Engeland, M., Roemen, G.M., Brink, M., Pachen Mm, W.M., De Bruïne Ap, *et al.* 2002. K-ras mutations and RASSF1A promoter methylation in colorectal cancer. *Oncogene,* 21**,** 3792-5.
Van Swearingen, A.E., Siegel, M.B. & Anders, C.K. 2014. Breast cancer brain metastases: evidence for neuronal-like adaptation in a 'breast-to-brain' transition? *Breast Cancer Res. ,* 16**,** 304.
Vanharanta, S. & Massague, J. 2013. Origins of metastatic traits. *Cancer Cell,* 24**,** 410-21.
Vanhoutteghem, A., Maciejewski-Duval, A., Bouche, C., Delhomme, B., *et al.* 2009. Basonuclin 2 has a function in the multiplication of embryonic craniofacial mesenchymal cells and is orthologous to disco proteins. *Proc Natl Acad Sci U S A,* 106**,** 14432-7.
Varela, I., Tarpey, P., Raine, K., Huang, D., *et al.* 2011a. Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. *Nature,* 484**,** 130.
Varela, I., Tarpey, P., Raine, K., Huang, D., *et al.* 2011b. Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. *Nature,* 469**,** 539-42.
Vasiliou, V., Pappa, A. & Petersen, D.R. 2000. Role of aldehyde dehydrogenases in endogenous and xenobiotic metabolism. *Chemico-Biological Interactions,* 129**,** 1-19.
Veeck, J., Noetzel, E., Bektas, N., Jost, E., *et al.* 2008. Promoter hypermethylation of the SFRP2 gene is a high-frequent alteration and tumor-specific epigenetic marker in human breast cancer. *Molecular cancer,* 7**,** 83.
Veeck, J., Wild, P.J., Fuchs, T., Schuffler, P.J., *et al.* 2009. Prognostic relevance of Wnt-inhibitory factor-1 (WIF1) and Dickkopf-3 (DKK3) promoter methylation in human breast cancer. *BMC cancer,* 9**,** 217.

Vinas-Castells, R., Beltran, M., Valls, G., Gomez, I.*, et al.* 2010. The hypoxia-controlled FBXL14 ubiquitin ligase targets SNAIL1 for proteasome degradation. *J Biol Chem,* 285**,** 3794-805.
Visvanathan, J., Lee, S., Lee, B., Lee, J.W. & Lee, S.K. 2007. The microRNA miR-124 antagonizes the anti-neural REST/SCP1 pathway during embryonic CNS development. *Genes Dev,* 21**,** 744-9.
Wagh, P.K., Peace, B.E. & Waltz, S.E. 2008. Met-Related Receptor Tyrosine Kinase Ron in Tumor Growth and Metastasis. 100**,** 1-33.
Wang, B., Zhao, B., North, P., Kong, A.*, et al.* 2013a. Expression of NgBR is highly associated with estrogen receptor alpha and survivin in breast cancer. *PLoS One,* 8**,** e78083.
Wang, D., Zhang, H., Li, M., Frid, M.G.*, et al.* 2014a. MicroRNA-124 controls the proliferative, migratory, and inflammatory phenotype of pulmonary vascular fibroblasts. *Circ Res,* 114**,** 67-78.
Wang, J., Zhang, S., Schultz, R.M. & Tseng, H. 2006. Search for basonuclin target genes. *Biochem Biophys Res Commun,* 348**,** 1261-71.
Wang, M.J., Li, Y., Wang, R., Wang, C.*, et al.* 2013b. Downregulation of microRNA-124 is an independent prognostic factor in patients with colorectal cancer. *Int J Colorectal Dis,* 28**,** 183-9.
Wang, P., Chen, L., Zhang, J., Chen, H.*, et al.* 2014b. Methylation-mediated silencing of the miR-124 genes facilitates pancreatic cancer progression and metastasis by targeting Rac1. *Oncogene,* 33**,** 514-24.
Wang, Y., Klijn, J.G.M., Zhang, Y., Sieuwerts, A.M.*, et al.* 2005. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *The Lancet,* 365**,** 671-9.
Wang, Y. & Shang, Y. 2013. Epigenetic control of epithelial-to-mesenchymal transition and cancer metastasis. *Exp Cell Res,* 319**,** 160-9.
Webster, K.A. 1987. Regulation of glycolytic enzyme RNA transcriptional rates by oxygen availability in skeletal muscle cells. *Mol Cell Biochem. ,* 77**,** 19-28.
Wei, J., Gabrusiewicz, K. & Heimberger, A. 2013. The controversial role of microglia in malignant gliomas. *Clin Dev Immunol. ,* 2013**,** 1-12.
Weidner, N., Semple, J.P., Welch, W.R. & Folkman, J. 1991. Tumor angiogenesis and metastasis--correlation in invasive breast carcinoma. *N Engl J Med.,* 324**,** 1-8.
Weigelt, B., Peterse, J.L. & Van 'T Veer, L.J. 2005. Breast cancer metastasis: markers and models. *Nat Rev Cancer,* 5**,** 591-602.
Weil, R.J., Palmieri, D.C., Bronder, J.L., Stark, A.M. & Steeg, P.S. 2005. Breast cancer metastasis to the central nervous system. *Am J Pathol,* 167**,** 913-20.
Weinhold, N., Jacobsen, A., Schultz, N., Sander, C. & Lee, W. 2014. Genome-wide analysis of noncoding regulatory mutations in cancer. *Nat Genet,* 46**,** 1160-5.
Weisenberger, D.J., Campan, M., Long, T.I., Kim, M.*, et al.* 2005. Analysis of repetitive element DNA methylation by MethyLight. *Nucleic Acids Research,* 33**,** 6823-36.
Westerveld, G.H., Repping, S., Leschot, N.J., Van Der Veen, F. & Lombardi, M.P. 2005. Mutations in the human BOULE gene are not a major cause of impaired spermatogenesis. *Fertil Steril,* 83**,** 513-5.
Willemen, H., Huo, X., Mao-Ying, Q. Zijlstra, J., Heijnen, C. And Kavelaars, A. 2012. MicroRNA-124 as a novel treatment for persistent hyperalgesia *Journal of Neuroinflammation,* 9**,** 2-10.
Willipinski-Stapelfeldt, B., Riethdorf, S., Assmann, V., Woelfle, U.*, et al.* 2005. Changes in cytoskeletal protein composition indicative of an epithelial-

mesenchymal transition in human micrometastatic and primary breast carcinoma cells. *Clin Cancer Res,* 11**,** 8006-14.
Wilting, S.M., Van Boerdonk, R.A., Henken, F.E., Meijer, C.J.*, et al.* 2010. Methylation-mediated silencing and tumour suppressive function of hsa-miR-124 in cervical cancer. *Mol Cancer,* 9**,** 167.
Wong, C., Lee, J.M., Ching, Y., Jin, D. & Oi-Lin, I. 2003. Genetic and epigenetic alterations of DLC1 gene in hepatocellualar carcinoma. *Cancer Research,* 63**,** 7646-51.
Wong, K.Y., So, C.C., Loong, F., Chung, L.P.*, et al.* 2011. Epigenetic inactivation of the miR-124-1 in haematological malignancies. *PLoS One,* 6**,** e19027.
Wu, X., Senechal, K., Mehran, S., Whang, Y.E. & Sawyers, C.L. 1998. The PTEN MMAC1 tumor suppressor phosphatase functions as a negative regulator of the phosphoinositide 3-kinase Akt pathway. *Proc. Natl. Acad. Sci,* 95**,** 15587-91.
Wu, Y., Zhang, L., Zhang, L., Wang, Y.*, et al.* 2015. Long non-coding RNA HOTAIR promotes tumor cell invasion and metastasis by recruiting EZH2 and repressing E cadherin in oral squamous cell carcinoma. *Int J Oncol.***,** 2586-94.
Xiao, W., Zhou, S., Xu, H., Li, H.*, et al.* 2013. Nogo-B promotes the epithelial-mesenchymal transition in HeLa cervical cancer cells via Fibulin-5. *Oncol Rep,* 29**,** 109-16.
Xu, G.L., Bestor, T.H., Bourc'his, D., Hsieh, C.L.*, et al.* 1999. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature,* 402**,** 187-91.
Yamada, K.M. & Araki, M. 2001. Tumor suppressor PTEN: modulator of cell signaling, growth, migration and apoptosis. *Signal Transduction and Cellular Organization,* 114**,** 2375-82.
Yamakuchi, M., Ferlito, M. & Lowenstein, C.J. 2008. miR-34a repression of SIRT1 regulates apoptosis. *Proc Natl Acad Sci U S A,* 105**,** 13421-6.
Yan, J., Yan, F., Li, Z., Sinnott, B.*, et al.* 2014. The 3M complex maintains microtubule and genome integrity. *Mol Cell,* 54**,** 791-804.
Yang, J. & Weinberg, R.A. 2008. Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. *Dev Cell,* 14**,** 818-29.
Yang, L., Chen, Y., Cui, T., Knosel, T.*, et al.* 2012. Desmoplakin acts as a tumor suppressor by inhibition of the Wnt/beta-catenin signaling pathway in human lung cancer. *Carcinogenesis,* 33**,** 1863-70.
Yang, Z., Gallicano, G.I., Yu, Q.C. & Fuchs, E. 1997. An unexpected localization of basonuclin in the centrosome, mitochondria, and acrosome of developing spermatids. *J Cell Biol.,* 137**,** 657-69.
Yi, J.M., Guzzetta, A.A., Bailey, V.J., Downing, S.R.*, et al.* 2013. Novel methylation biomarker panel for the early detection of pancreatic cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research,* 19**,** 6544-55.
Yoo, C.B. & Jones, P.A. 2006. Epigenetic therapy of cancer: past, present and future. . *Nat Rev Drug Discov.,* 5**,** 121.
Yoon, H.-G., Chan, D.W., Reynolds, A.B., Qin, J. & Wong, J. 2003. N-CoR Mediates DNA Methylation-Dependent Repression through a Methyl CpG Binding Protein Kaiso. *Molecular Cell,* 12**,** 723-34.
Yuan, J., Luo, R.Z., Fujii, S., Wang, L.*, et al.* 2003. Aberrant Methylation and Silencing of ARHI, an Imprinted Tumor Suppressor Gene in which the Function Is Lost in Breast Cancers. *Cancer Res,* 63**,** 4174-80.

Zaman, M.S., Thamminana, S., Shahryari, V., Chiyomaru, T.*, et al.* 2012. Inhibition of PTEN gene expression by oncogenic miR-23b-3p in renal cancer. *PLoS One,* 7**,** e50203.
Zeng, H., Irwin, M.L., Lu, L., Risch, H.*, et al.* 2012. Physical activity and breast cancer survival: an epigenetic link through reduced methylation of a tumor suppressor gene L3MBTL1. *Breast Cancer Res Treat,* 133**,** 127-35.
Zetter, B.R. 1998. Angiogenesis and tumor metastasis. *Annu Rev Med.,* 49**,** 407-24.
Zhang, C. & Yu, D. 2011. Microenvironment determinants of brain metastasis. *Cell Biosci,* 1**,** 8.
Zhang, S., Wang, J. & Tseng, H. 2007. Basonuclin regulates a subset of ribosomal RNA genes in HaCaT cells. *PLoS One,* 2**,** e902.
Zhang, X., Mu, W. & Zhang, W. 2012. On the analysis of the illumina 450k array data: probes ambiguously mapped to the human genome. *Front Genet,* 3**,** 73.
Zhang, X.H., Giuliano, M., Trivedi, M.V., Schiff, R. & Osborne, C.K. 2013. Metastasis dormancy in estrogen receptor-positive breast cancer. *Clin Cancer Res,* 19**,** 6389-97.
Zhang, Y., Iwasaki, H., Wang, H., Kudo, T.*, et al.* 2003. Cloning and characterization of a new human UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase, designated pp-GalNAc-T13, that is specifically expressed in neurons and synthesizes GalNAc alpha-serine/threonine antigen. *J Biol Chem,* 278**,** 573-84.
Zhang, Y.N.H., Erdjument-Bromage, H., Tempst, P., Bird, A. & Reinberg, D. 1999. Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev.,* 13**,** 1924-35.
Zheng, H., Du, Y., Hua, Y., Wu, Z.*, et al.* 2012. Essential role of Fbxl14 ubiquitin ligase in regulation of vertebrate axis formation through modulating Mkp3 level. *Cell Res,* 22**,** 936-40.
Zhong, H., De Marzo, A.M., Laughner, E., Lim, M.*, et al.* 1999. Overexpression of Hypoxia-inducible Factor 1 α in Common Human Cancers and Their Metastases. *Cancer Cell,* 59**,** 5830-8.

Publications

Peer reviewed articles

1. **Rajendra P Pangeni**, Prasanna Channathodiyil David Huen, Balraj K Johal, Dawar Pasha, Lawrence W Eagles, Natasha Hadjistephanou, Preethi Chandrasekaran, Oliver Nevell, Claire L Davies, Ayobami I Adewumi, Hamida Khanom, Ikroop S Samra, Caitlin Sankey, Sophie Willetts, Vannessa C. Buzatto, Thoraia Shinawi, Timothy P Dawson, Katherine M Ashton, Charles Davis, Carol Walker, Michael Jenkinson, Ivan Bièche, Farida Latif, John L Darling, Tracy J Warr, Mark R Morris (2015). The *GALNT9*, *BNC1* and *CCDC8* genes are frequently epigenetically dysregulated in breast tumours that metastasise to the brain. ***Clinical Epigenetics***. **7** (1). 57-74

2. **Rajendra P Pangeni**, David Huen, Vannessa C. Buzatto, Timothy P Dawson, Katherine M Ashton, Charles Davis, Carol Walker, Michael Jenkinson, Ivan Bièche, Farida Latif, John L Darling, Tracy J Warr , Mark R Morris (2015). Genome Wide methylation analyses identify novel genes including microRNAs and non-protein coding genes as prognostic biomarkers in breast cancer metastases to the brain. **Manuscript in preparation**.

Conference papers and published abstracts

1 **Rajendra P Pangeni**, David Huen, Katherine Ashton, Carol Walker, Timothy P Dawson, Charles H G Davis, Farida Latif, John L Darling, Tracy J Warr, Mark R Morris (2014). Multi-gene methylation analysis to identify signature genes for brain metastasis from primary breast tumors. ***Neuro Oncol***, **16** (Suppl 5), v37. ABSTRACT. Paper presented in Society of Neuro Oncology (SNO) conference at Miami, United States, 13th to 16th November 2014.

2 **Rajendra P Pangeni**, David Huen, Katherine Ashton, Carol Walker, Timothy P Dawson, Charles H G Davis, Farida Latif, John L Darling, Tracy J Warr, Mark R Morris (2014). Identifying epigenetic changes in breast tumours that

metastasise to the brain. ABSTRACT. Paper presented in National Cancer Research Institute (NCRI) conference at Liverpool, 2nd to 5th November 2014

3 **Rajendra P Pangeni,** Kate Ashton, Carol Walker, Tim Dawson, Charles Davis, Farida Latif, John L Darling, Tracy J Warr, Mark R Morris (2014), Identification of genes epigenetically deregulated in brain metastasis from primary breast tumours. ***Neuro Oncol***, **16** (Suppl 6), vi18. ABSTRACT. Paper presented in British Neuro Oncology Society (BNOS) at Liverpool, 9th to 11th July 2014.

4 **Rajendra P Pangeni,** Kate Ashton, Carol Walker, Tim Dawson, Charles Davis, Farida Latif, John L Darling, Tracy J Warr, Mark R Morris (2013). Multi-gene methylation analysis to identify signature genes for brain metastasis from primary breast tumours. ABSTRACT. Paper presented in British Neuro Oncology Society (BNOS) at Durham. 10th to 12th July 2013).

5 **Rajendra P Pangeni,** Tracy J Warr, Mark R Morris (2012). Identifying epigenetic determinant of metastatic brain tumours. ***Neuro Oncol***, **14** (Suppl 2), ii1-ii12. ABSTRACT. Paper presented in British Neuro Oncology Society (BNOS) at Manchester. 26th to 28th June 2012**.**

Pangeni et al. Clinical Epigenetics (2015) 7:57
DOI 10.1186/s13148-015-0089-x

CLINICAL EPIGENETICS

RESEARCH Open Access

The *GALNT9*, *BNC1* and *CCDC8* genes are frequently epigenetically dysregulated in breast tumours that metastasise to the brain

Rajendra P. Pangeni[1], Prasanna Channathodiyil[1], David S. Huen[2], Lawrence W. Eagles[1], Balraj K. Johal[2], Dawar Pasha[2], Natasa Hadjistephanou[2], Oliver Nevell[2], Claire L. Davies[2], Ayobami I. Adewumi[2], Hamida Khanom[2], Ikroop S. Samra[2], Vanessa C. Buzatto[2], Preethi Chandrasekaran[2], Thoraia Shinawi[3], Timothy P. Dawson[4], Katherine M. Ashton[4], Charles Davis[4], Andrew R. Brodbelt[5], Michael D. Jenkinson[5], Ivan Bièche[6], Farida Latif[3], John L. Darling[1], Tracy J. Warr[1] and Mark R. Morris[1,2,3*]

Abstract

Background: Tumour metastasis to the brain is a common and deadly development in certain cancers; 18–30 % of breast tumours metastasise to the brain. The contribution that gene silencing through epigenetic mechanisms plays in these metastatic tumours is not well understood.

Results: We have carried out a bioinformatic screen of genome-wide breast tumour methylation data available at The Cancer Genome Atlas (TCGA) and a broad literature review to identify candidate genes that may contribute to breast to brain metastasis (BBM). This analysis identified 82 candidates. We investigated the methylation status of these genes using Combined Bisulfite and Restriction Analysis (CoBRA) and identified 21 genes frequently methylated in BBM. We have identified three genes, *GALNT9*, *CCDC8* and *BNC1*, that were frequently methylated (55, 73 and 71 %, respectively) and silenced in BBM and infrequently methylated in primary breast tumours. *CCDC8* was commonly methylated in brain metastases and their associated primary tumours whereas *GALNT9* and *BNC1* were methylated and silenced only in brain metastases, but not in the associated primary breast tumours from individual patients. This suggests differing roles for these genes in the evolution of metastatic tumours; *CCDC8* methylation occurs at an early stage of metastatic evolution whereas methylation of *GANLT9* and *BNC1* occurs at a later stage of tumour evolution. Knockdown of these genes by RNAi resulted in a significant increase in the migratory and invasive potential of breast cancer cell lines.

Conclusions: These findings indicate that *GALNT9* (an initiator of O-glycosylation), *CCDC8* (a regulator of microtubule dynamics) and *BNC1* (a transcription factor with a broad range of targets) may play a role in the progression of primary breast tumours to brain metastases. These genes may be useful as prognostic markers and their products may provide novel therapeutic targets.

Keywords: Breast, Brain, Metastasis, DNA methylation, Epigenetic, Tumour suppressor

* Correspondence: morrismr@wlv.ac.uk
[1]Brain Tumour Research Centre, University of Wolverhampton, Wolverhampton, UK
[2]School of Biology, Chemistry and Forensic Sciences, University of Wolverhampton, Wolverhampton, UK
Full list of author information is available at the end of the article

Background

Brain metastases account for up to 40 % of all secondary tumours, with an estimated 27,000 new cases every year in the UK [1, 2]. Current estimates suggest that 18–30 % of patients with breast cancer eventually develop brain metastases [3–6]. The frequency of metastatic brain tumours is rising; this increased incidence is due, in part, to an ageing population, improved neuroimaging and increased patient survival following primary tumour treatment [7]. Currently, brain metastases are treated by whole brain radiotherapy, stereotactic radiosurgery and surgical resection either individually or in combination [8]. However, following treatment, patient prognosis remains poor; both morbidity and mortality are high and the median survival is approximately 7 months [9].

Evidence indicating that tumours originating in specific organs favour certain sites of metastasis has existed for over 50 years [10]. However, the underlying mechanisms of this organotropism towards specific secondary sites such as the brain are still poorly understood. Although the genetic basis of primary tumour formation is becoming increasingly clear [11], it is still unclear which of the many hundreds of tumour-associated alterations found in primary breast cancer [12, 13] contribute to metastasis and moreover, metastasis to specific secondary sites such as the brain. The primary tumour types that most frequently metastasise to the brain are lung, breast, melanoma and renal cancers. However, the speed at which these secondary tumours develop varies greatly with breast to brain metastases (BBM) occurring relatively slowly [7]. This specificity indicates that, at least in part, genomic alterations that drive tumour formation in these primary organs provide the potential for colonization of a distinct subset of secondary organ sites.

There is little in the way of prognostic markers for BBM. It is known that the risk of BBM occurring early (<2 years after primary diagnosis) is associated with early onset tumours, estrogen receptor negative (ER-ve), human epidermal growth factor receptor 2 overexpression (HER2 + ve) and triple negative (ER-ve/PR-ve/HER2-ve) phenotypes [14–17]. However, more than 50 % of BBMs occur over 5 years after the primary tumour was diagnosed. Many of these late recurring brain metastases are derived from ER+ primary tumours [4, 9, 18]. The common long lag-time between primary tumour diagnosis and recurrence of a detectable secondary tumour suggests that cells from these breast tumours undergo a period of dormancy [19, 20]. These dormant cells are often found as micrometastases in bone marrow. However, the presence of these micrometastases is not in itself a strong prognostic indicator for later metastatic disease [21, 22]. It is possible that brain micrometastases are common, and these require further genomic alterations to occur before sustained proliferation and growth occurs.

Genomic alterations that provide the potential for metastatic growth can be characterised as either those that also drive primary tumour growth advantage, those that provide potential for dissemination and infiltration [23] or those that enable continued growth within the microenvironment of the new organ [24]. A number of genetic and epigenetic alterations acquired by breast tumour micrometastases of the bone have been characterised [25, 26]. However, very little is known about specific genomic alterations that facilitate colonisation in the brain.

We have carried out a screen to identify genes frequently dysregulated through promoter hypermethylation in BBM. This analysis has identified candidate genes that are either dysregulated early in tumour evolution (methylation is common to primary tumour and resulting BBM) or at a later stage, once the cells that will evolve into the BBM have disseminated from the primary tumour. We hope that this preliminary analysis may provide initial evidence of novel targets that can be utilised in the development of prognostic screens and new rational therapeutic approaches for breast tumours and brain metastases.

Methods

Selection of candidate metastatic suppressor genes

For an overview of our candidate selection strategies see Additional file 1: Figure S1. We utilized the Illumina HumanMethylation 450 K BeadChip methylation array data from *The Cancer Genome Atlas* (TCGA) to identify candidate genes (Additional file 2: Table S1 for TCGA tumour barcodes). To ensure we were selecting genuine promoter-associated CpG islands, we selected only those probes that are located in the 5′ region of the gene or up to 1500 base pairs from the transcription start site (identified in the array annotation as TSS, TSS200, TSS1500). We identified individual probes that are not methylated (β value ≤ 0.25) in 75 % (15/20) of primary breast tumours and methylated (β value ≥ 0.60) in primary lung tumours, in at least 50 % (10/20) of the samples. This analysis generated four candidates that were then characterised in the laboratory.

In addition to our bioinformatic analysis, we carried out a broad literature review to identify candidate genes. We generated a long-list of genes that had previously been identified as hypermethylated in one of the primary tumours types that readily metastasise to the brain (lung, melanoma or renal [7]). We expanded this long-list by selecting genes that are downregulated in epithelial-mesenchymal transition (EMT) and that possess a well-defined promoter region CpG island. By interrogating all the available breast tumour methylation data in the TCGA by using their data portal (https://tcga-data.nci.nih.gov/tcga/), we shortlisted only those genes that were infrequently methylated in primary

breast tumours. This analysis generated 78 candidates that were then characterised in the laboratory.

Patients and samples

Thirty-one fresh-frozen metastatic brain tumours originating from primary breast tumours were provided by The Walton Research Tissue Bank, Liverpool and Brain Tumour North West (BTNW) Tissue Bank, Preston. Eleven pairs of formalin fixed paraffin embedded (FFPE) primary breast tumours corresponding to matched metastatic brain tumours were provided by BTNW tissue bank. Receptor status information is available for 9 of the 11 primary tumour pairs, six of these are ER + ve, one is triple negative. The time between primary tumour surgery and removal of the brain metastasis ranges from 2 to 10 years (Additional file 3: Table S2a).

A cohort of 40 independent primary breast tumours was also analysed. All breast tumours from this cohort were ducal carcinomas; their clinical characteristics are described in [27]. Molecular characterisation was available for 20 of these tumours, 15 of these are ER + ve and three are triple negative. No brain metastases were observed in any of these patients, nine patients had been screened for metastasis 10 years or more post-primary tumour surgery and 17 of the 20 after more than 5 years (Additional file 3: Table S2b).

Tissues were obtained from National research Ethics committee approved research tissue banks, and informed consent was obtained from each patient. This study was conducted according to the principles expressed in the Declaration of Helsinki.

Breast cancer cell Lines and 5-Aza-2′–deoxycytidine treatment

Five breast cancer cell lines (MCF7, T74D, MDA-MB231, BT549 and ZR75) were routinely maintained in DMEM (Sigma, UK) supplemented with 10 % FCS at 37 °C and 5 % CO_2. Cells were plated according to their doubling time to ensure that both control and 5-AZA-2′-deoxycytidine (5-AZA-dC; Sigma, UK)-treated cells lines were approximately 75 % confluent at the time of RNA extraction. 5-AZA-dC was freshly prepared in ddH_2O and filter sterilized. Twenty-four hours after seeding, cells were treated with 5 μM 5-AZA-dC. Cells were treated with fresh 5 μM 5-AZA-dC three times a week on alternate days. After 7 days, the cells were collected using 1 % trypsin; cell pellets were washed with PBS.

Genomic DNA/RNA extraction

Genomic DNA was extracted from fresh-frozen metastatic brain tumours using The *DNA isolation kit from cells and tissues* (Roche, Germany). Briefly, 25 mg of tissue was homogenised using lysis buffer and incubated at 37 °C for 30 min followed by addition of Proteinase K and RNase solution. The samples were then centrifuged and processed according to manufacturer's instructions. For FFPE samples, a *FFPE DNA extraction kit* (Qiagen, USA) was used. Briefly, a small block of samples embedded with paraffin was cut into thin sections and mixed with xylene followed by 100 % ethanol. The samples were then processed according to manufacturer's instructions. Similarly, total RNA was extracted using the *EZ-RNA extraction kit* (Biological Industries, Israel). Briefly, fresh-frozen tumours were homogenized using lysis buffer followed by addition of extraction solution. The samples were then centrifuged and processed according to manufacturer's instructions. DNA concentration was measured using a *nanodrop2000* (Thermo Scientific, USA).

Bisulfite conversion of DNA

Bisulfite conversion of genomic DNA from metastatic brain tumours (500 ng) and positive controls was carried out using the *EZ DNA methylation kit* (Zymo Research Corp., USA) according to manufacturer's instruction. Fully methylated, positive controls were generated by incubating gDNA with DNA methyltransferase, in the presence of S-Adenosyl methionine (SAM) (New England bio lab, USA) for 2 h at 37 °C prior to bisulfite conversion.

Promoter methylation analysis

Primers used to amplify promoter regions from bisulfite-modified DNA can be found in Additional file 4: Table S3. Primers were designed based on standard bisulfite DNA primer designing criteria [28]. These primers were used to amplify bisulfite converted DNA. DNA methylation was determined by digesting Combined Bisulfite and Restriction Analysis (CoBRA) PCR products with the BstUI and TaqI restriction enzymes (Fementas, UK).

Quantitative methylation analysis of tumour DNA was carried out by cloning bisulfite-PCR products (individual alleles) into pGEM plasmid (Promega, UK) followed by sequencing of individual clones using primers to M13.

The CpG island regions of *BNC1*, *CCDC8* and *GALNT9* are presented in Additional file 5: Figure S2, details of PCR primer sites and individual CpG dinucleotides analysed by sequencing are provided.

Migration assay

Candidate genes were knocked down in breast cancer cell lines by transfection of RNAi 'silencer select' oligos against *CCDC8* (s228331), *BNC1* (s2012) or *GALNT9* (s27040), control cells were transfected with control oligo no. 1 (Ambion, Austin, TX, USA). After 24 h, DMEM with 10 % FBS was replaced with fresh DMEM without FBS and incubated at 37 °C for 24 h. Confluent monolayer of cells in each well was scratched with the tip of a 200 μl pipette tip. The extent of migration of cells was observed after 24 and 48 h.

Invasion assay

Two hundred microlitres of matrigel matrix (Becton Dickinson, NJ, USA) was applied to 24-well 9-mm inserts containing polyethylene terephthalate (PET) membranes with 8-um pores (Corning, USA). One hundred fifty thousand cells were applied to the invasion chamber. DMEM containing 10 % FBS was placed in the lower chamber as a chemoattractant. The plates were incubated at 37 °C for 48 h with 5 % CO_2. Cells from the lower layer were stained with crystal violet. The optical density at 540 nm for each well was determined.

Western blotting

Cells were lysed in RIPA buffer (25 mM HCL, 0.1 % SDS, 1 % triton 100, 0.15 M NaCl) containing phosphatase and protease inhibitor (Roche, Germany). Seventy micrograms of each extract was resolved on polyacrylamide gels and probed with anti-rabbit primary antibody against *CCDC8* (ab170233), *BNC1* (ab123645) or *GALNT9* (ab173682) (Abcam, USA). Signals were detected with horseradish peroxidase-conjugated anti-rabbit antibody (GE Healthcare, UK) and enhanced chemiluminescence (Biological Industries, Israel). Membranes were stained with India ink (Winsor and Newton, UK) for comparison of loading.

Results

Screening of candidate BBM suppressor genes

We have used a candidate gene approach to identify genes deregulated in breast tumours that metastasise to the brain. See Additional file 6: Table S4 for details of genes.

Bioinformatic analysis of primary tumour genome-wide methylation arrays identified four candidate BBM suppressor genes

We have compared the methylation status (β value) of array probes in *TCGA* data sets from 20 primary breast tumours (with no evidence of metastatic disease) and 20 primary lung tumours. We hypothesised that genes that are infrequently methylated in non-metastasising breast tumours and frequently methylated in primary lung tumours that readily metastasis to the brain (metastases are identified relatively soon after primary tumour diagnosis) [29] may be commonly methylated in metastatic brain tumours that derive from both lung and breast tumours.

We filtered probes in primary breast tumours to identify those that are infrequently methylated, (having a β value ≤ 0.25 in at least 15/20 (75 %)). This resulted in 97,155 probes. Filtering of frequently methylated probes, (having a β value ≥ 0.60 in at least 50 % (10/20)) in lung tumours resulted in 45,382 probes. Comparison of the probes between breast and lung tumours identified eight probes that corresponded to six genes (*GALNT9*, *KRT222*, *PLEKHA6*, *TFAP2A*, *TSPAN4* and *ZNF808*). Two of these genes (*KRT222* and *PLEKHA6*) do not have well-defined CpG islands. In total, this genome wide approach identified four candidate genes (*GALNT9*, *TFAP2A*, *TSPAN4* and *ZNF808*) for further analysis.

A literature review identified 78 candidates BBM suppressor genes

We have screened genes that have previously been shown to be frequently methylated and silenced in at least one of the primary tumours types that rapidly (relative to many breast tumours) metastasise to the brain, i.e. lung, melanoma and renal cell carcinoma (RCC) [7]. We then interrogated *TCGA* to determine the methylation status of these genes in primary breast tumours. This screen identified 42 candidate metastatic suppressor genes that are infrequently methylated in primary breast cancer and frequently methylated in primary lung, melanoma or renal tumours (Additional file 6: Table S4 and references therein).

In addition, we selected 36 metastasis suppressor candidates that are downregulated during EMT (Additional file 6: Table S4 and references therein).

Identification of frequently methylated genes in metastatic brain tumours

The methylation status of 82 candidate genes was determined by CoBRA [28] in 30 BBM. To ensure that we were identifying genes which are enriched in the population of patients with BBM that are most likely to be clinically significant, we have imposed a high cut-off of ≥50 % of all metastatic tumours being methylated for a gene to be considered as frequently methylated. For this preliminary screening, we have determined that a significant proportion of the promoters within the tumour sample is methylated if there are clearly observed digest products following restriction analysis.

From the panel of four genes selected from our analyses of HumanMethylation 450 K BeadChip arrays obtained from TCGA, only one gene, *GALNT9*, was frequently methylated (55 %) in metastatic brain tumours originating from primary breast tumours (see Table 1, Fig. 1a, Additional file 7: Figure S3).

From our panel of 42 literature review candidate genes, we identified ten genes that are frequently methylated in brain metastases. These were *HOXD3* (100 %), *CCDC8* (73 %), *HOXB13* (80 %), *ABCB1* (80 %), *PENK* (80 %), *BNC1* (71 %), *PCDH8* (53 %), *STAT3* (67 %), *TNFRSF10D* (60 %) and *WIF1* (53 %) (see Table 1, Fig. 1a, Additional file 7: Figure S3).

We proceeded to determine the methylation status of these ten genes in an independent cohort of primary breast tumours.

Table 1 Genes frequently methylated in breast to brain metastases. Twenty-one genes are frequently methylated in brain metastases ($n = 15$) of which 18 genes are also frequently methylated in primary breast tumours ($n = 20$). Three genes, *CCDC8, BNC1* and *GALNT9* (highlighted in grey), are infrequently methylated in primary breast tumours. These genes were further analysed in 20 primary breast samples ($n = 40$ in total) and 15 breast to brain metastases ($n = 30$ in total)

Gene symbol	Accession	Gene name	% of metastatic tumours methylated	Function
CLDN18	NM_016369.3	*Claudin 18*	100	Intercellular adhesion molecule responsible for tight junction strand formation [77]
KRT85	NM_002283.3	*Keratin 85*	100	Component of intermediate filament in epithelial cells contributing to cell-cell adhesion [78–80]
MIR127	NR_029696.1	*microRNA 127*	100	Regulator of cell proliferation and senescence [81]
MIR433	NR_029966.1	*microRNA 433*	100	Deregulated in gastric cancer, regulator of cell migration and drug response [82, 83]
HOXD3	NM_006898.4	*HomeoboxD3*	100	Proangiogenic transcription factor [84]
MIR23B	NR_029664.1	*microRNA 23b*	92	Involved in cytoskeleton modelling, motility and metastasis [85–88]
CCDC8	NM_032040.4	*Coil coiled domain containing 8*	73	Mutated in patients with 3 M syndrome [70]. Loss is associated with genomic instability and aneuploidy [75].
KRT83	NM_002282.3	*Keratin 83*	84	Component of intermediate filament, contributes to cell to cell adhesion [78, 80]
HOXB13	NM_006361.5	*Homeobox B13*	80	TSG for prostate cancer, inhibits androgen mediated signalling [89]
ABCB1	NM_000927.4	*ATP-binding cassette sub-family B member 1*	80	Controls efflux of substances across plasma membranes, associated with multidrug resistance [90]
PENK	NM_006211.3	*Proenkephalin*	80	Promotes RNA splicing in osteoblasts and neural cells, plays role in bone development [91]
MST1R	NM_002447.2	*Macrophage stimulating 1 receptor*	78	Involved in intracellular signalling cascades leading to cellular growth, motility and invasion [92]
BNC1	NM_001717.3	*Basonuclin 1*	71	Zink finger transcription factor, regulator of EMT [68]
PCDH8	NM_002590.3	*Procadhern 8*	73	Helps in cell to cell adhesion [93]
STAT3	NM_139276.2	*Signal transducer and activator of transcription 3*	67	Involved in embryonic stem cell regulation, somatic cell growth [94–96]
BVES	NM_007073.4	*Blood vessel epicardial substance*	64	Involved in inter-cellular interaction and cell adhesion [97]
TNFRSF10D	NM_003840.4	*Tumour Necrosis Factor receptor superfamily 10 D*	60	Member of TNF (Tumour Necrosis Factor) receptor superfamily, promotes apoptosis in cancer cells [98]
CLDN6	NM_021195.4	*Claudin 6*	55	Intercellular adhesion molecules responsible for tight junction strand formation, its epigenetic silencing is associated with migration and invasiveness of breast cancer [77, 99]
HOXD10	NM_002148.3	*Homeobox D10*	55	Maintain epithelial cell plasticity and contributes to stability of extracellular matrix [100]
GALNT9	NM_001122636.1	*N-acetyl galactosaminyl transferase 9*	55	Catalyzes O-glycosylation [53, 101]
WIF1	NM_007191.4	*Wnt inhibitory factor-1 gene*	53	Inhibitor of Wnt-signalling [102, 103]

In addition, from a panel of 36 genes downregulated in EMT, we identified 10 genes frequently methylated in metastatic brain tumours originated from primary breast tumours. These were *CLDN18* (100 %), *KRT85* (100 %), *MIR127* (100 %), *MIR433* (100 %), *MIR23b* (92 %), *KRT83* (84 %), *MST1R* (78 %), *BVES* (64 %), *CLDN6* (55 %) and *HOXD10* (55 %) (see Table 1, Fig. 1a). We proceeded to determine the methylation status of these ten genes in an independent cohort of primary breast tumours.

A graphical overview of our candidate selection process and results is presented in Additional file 1: Figure S1.

GALNT9, BNC1 and *CCDC8* are differentially methylated in primary breast tumours and BBM

We have screened primary breast tumours for the presence of methylation in the genes that are frequently methylated in BBM. To ensure that genes identified in this study are clinically significant, we have imposed a

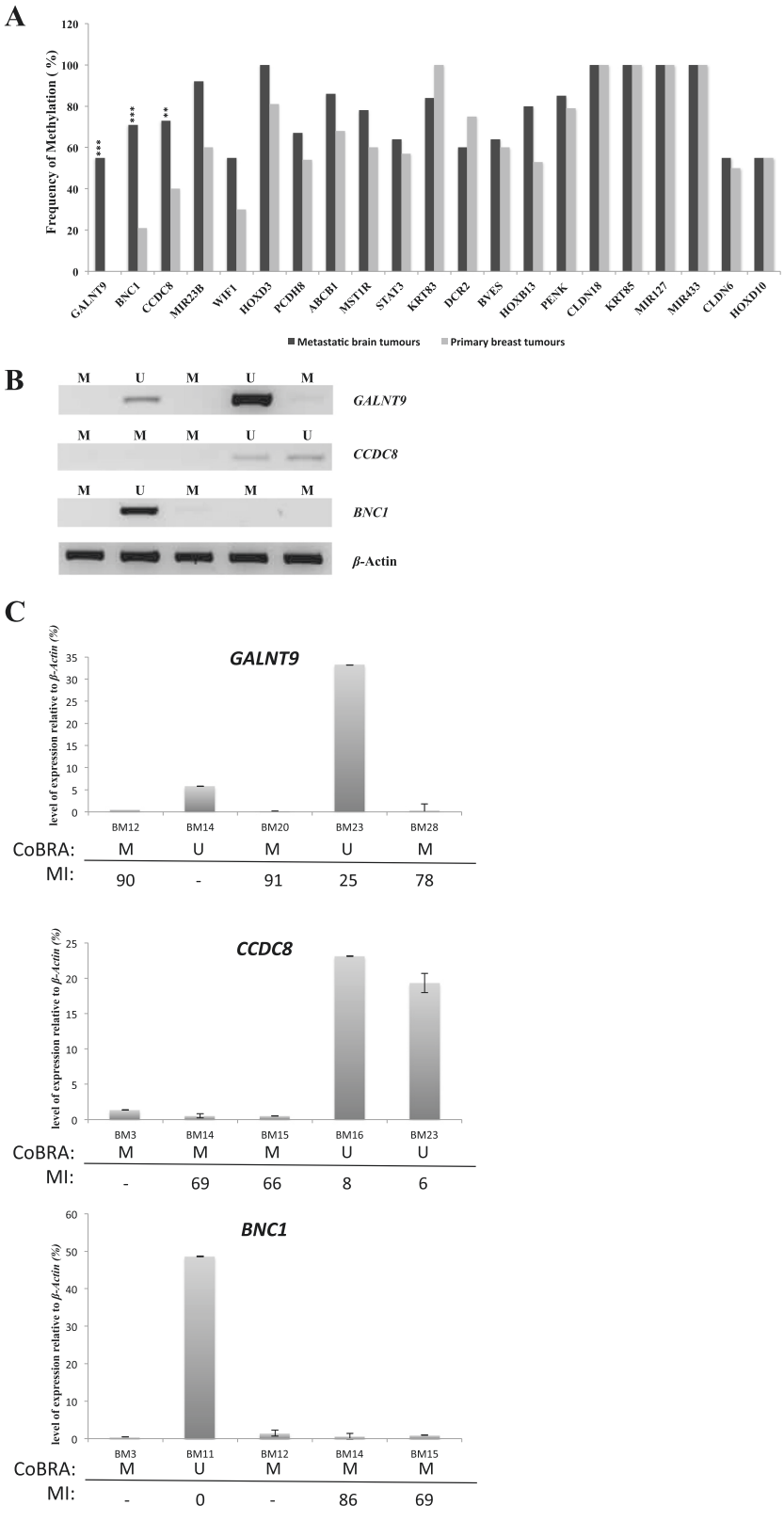


Fig. 1 (See legend on next page.)

(See figure on previous page.)
Fig. 1 a Methylation frequency of candidate metastatic suppressor genes in breast-to-brain metastases (BBM) ($n = 30$) versus primary breast tumours ($n = 40$). Out of the 25 genes that were frequently methylated in brain metastases, three genes (*GALNT9*, *CCDC8* and *BNC1*) were infrequently methylated in a cohort of independent primary tumours with statistical significance ($p = 0.0001$, 0.01 and 0.0001, respectively). **b**, **c** Expression of *GALNT9*, *CCDC8* and *BNC1* correlates with promoter methylation in BBM. Reverse transcription PCR (RT-PCR) for *GALNT9*, *CCDC8* and *BNC1* in BBM shows that these genes were expressed in tumours where their promoters are unmethylated (U) and silenced in methylated (M) tumours (see Additional file 7: Figs S3 and Additional file 9: Figure S5 for representative methylation analysis). Expression of β-actin was determined to ensure equal loading for all samples. **c** Expression levels of each gene were quantified in relation to the expression of β-actin. The methylation status was determined by either CoBRA or sequencing of individual alleles to determine the methylation index (MI) for individual tumours. High levels of expression were not associated with high levels of methylation in the region analysed. A full set of methylation/expression analysis can be seen in Additional file 10: Figure S6 (*BM* brain metastasis, *M* methylated, *U* unmethylated, – analysis was not done)

relatively low cut-off frequency of ≤45 % for methylation in primary breast tumours.

We analysed the 21 genes that were frequently methylated in BBM in a cohort of 40 primary breast tumours (unrelated to the brain metastasis cohort [27]).

We found *GALNT9* to be frequently methylated in BBM (55 %) and not methylated in any of the 40 primary breast tumours ($p = 0.0001$).

From a panel of ten genes frequently methylated in brain metastases (from our literature review candidates), we identified that eight of these genes were also frequently methylated in primary breast tumours. These are *HOXD3* (81 %), *HOXB13* (53 %), *ABCB1* (68 %), *PCDH8* (54 %), *PENK* (79 %), *STAT3* (57 %), *TNFRSF10D* (75 %) and *WIFI* (55 %) (Fig. 1a). This suggests that these genes are not uniquely epigenetically deregulated during the process of BBM. However, it is worth noting that to our knowledge this is the first time that promoter methylation in *CCDC8*, *HOXD3*, *PCDH8*, *PENK*, *STAT3*, *SFRP2* and *WIFI* has been described in primary breast tumours.

Promoter methylation of *BNC1* (17 %) and *CCDC8 (*40 %) in primary breast tumours was infrequent (≤45 %), and statistically significantly lower than that of the frequency of methylation in BBM ($p = 0.0001$ and 0.01, respectively) (Fig. 1a). The low frequency of methylation in primary tumours indicates that *BNC1* and *CCDC8* may contribute to BBM and are good candidates for further investigation.

We found that all ten EMT-related genes were frequently methylated in primary breast tumours, i.e. *CLDN18* (100 %), *KRT85* (100 %), *MIR127* (100 %), *MIR433* (100 %), *MIR23b* (60 %), *KRT83* (100 %), *MST1R* (60 %), *BVES* (60 %), *CLDN6* (50 %) and *HOXD10* (55 %) (Fig. 1a). The high frequency of methylation in primary tumours indicates that epigenetic deregulation of these genes is not driving BBM.

From our broad-ranging screens, we have identified *GALNT9*, *BNC1* and *CCDC8* as frequently methylated in BBM and significantly less frequently methylated in primary breast tumours (Fig. 1, Additional file 8: Figure S4).

To ensure that CoBRA digests were representative of high methylation status in tumours, we carried out base-resolution analysis of promoter region methylation for *BNC1*, *CCDC8* and *GALNT9* by cloning and sequencing individual bisulfite-modified alleles from select tumours (Additional file 9: Figure S5). This analysis was used to determine the methylation index (MI) of CpG islands for individual tumours. MI is defined as the total number of methylated CpG dinucleotides given as a percentage of all CpGs analysed. The MI for regions determined to be methylated by CoBRA ranged from 60 to 91 % whereas those promoters deemed not to be methylated by CoBRA had MIs ranging between 0 and 36 %. From this analysis, we have defined that, for these samples, physiologically significant methylation levels are those of ≥60 % MI and lack of physiologically significant methylation is defined as <40 % MI.

Expression analysis of *BNC1*, *CCDC8* and *GALNT9* in metastatic brain tumours

Having identified three candidate genes that are differentially methylated in primary breast tumours and metastatic brain tumours, we proceeded to determine if this promoter methylation correlated to gene expression.

Total RNA was extracted from 15 metastatic brain tumours to determine the expression of *BNC1*, *CCDC8* and *GALNT9* by RT-PCR. The expression level of each gene was quantified in relation to the expression of β-actin, in tumours with unmethylated promoters (MI = 0–25 %). The maximum expression of these genes was 49, 23 and 33 % that of β-actin, respectively. *BNC1*, *CCDC8* and *GALNT9* were frequently downregulated or silenced in these tumours and reduced expression correlated to promoter methylation as determined by CoBRA and base-resolution sequencing (Fig. 1b and c, Additional file 10: Figure S6). These genes were also commonly silenced in breast cancer cell lines, this silencing was reversed following treatment with 5-Aza-2′-deoxycytidine an inhibitor of DNA methyltransferase enzymes [30] (Additional file 11: Figure S7).

Promoter methylation status of *BNC1*, *CCDC8* and *GALNT9* in brain metastases and associated primary breast tumours from individual patients

We analysed the methylation status *BNC1*, *CCDC8* and *GALNT9* in metastatic brain tumours and corresponding

primary tumours from individual patients. We had ten pairs, however, some loci in the primary tumour DNA proved refractive to amplification. Of eight matched pairs, where the *BNC1* promoter region was successfully amplified, the region was methylated in all eight of the brain metastases. However, it was only methylated in one corresponding primary tumour (Fig. 2a). The *GALNT9* promoter was methylated in 3/5 brain metastases and not methylated in any of the corresponding primary breast tumours (Fig. 2b). In contrast, out of 11 matched pairs, *CCDC8* was commonly methylated in 10 corresponding primary tumours (Fig. 2c). This common *CCDC8*

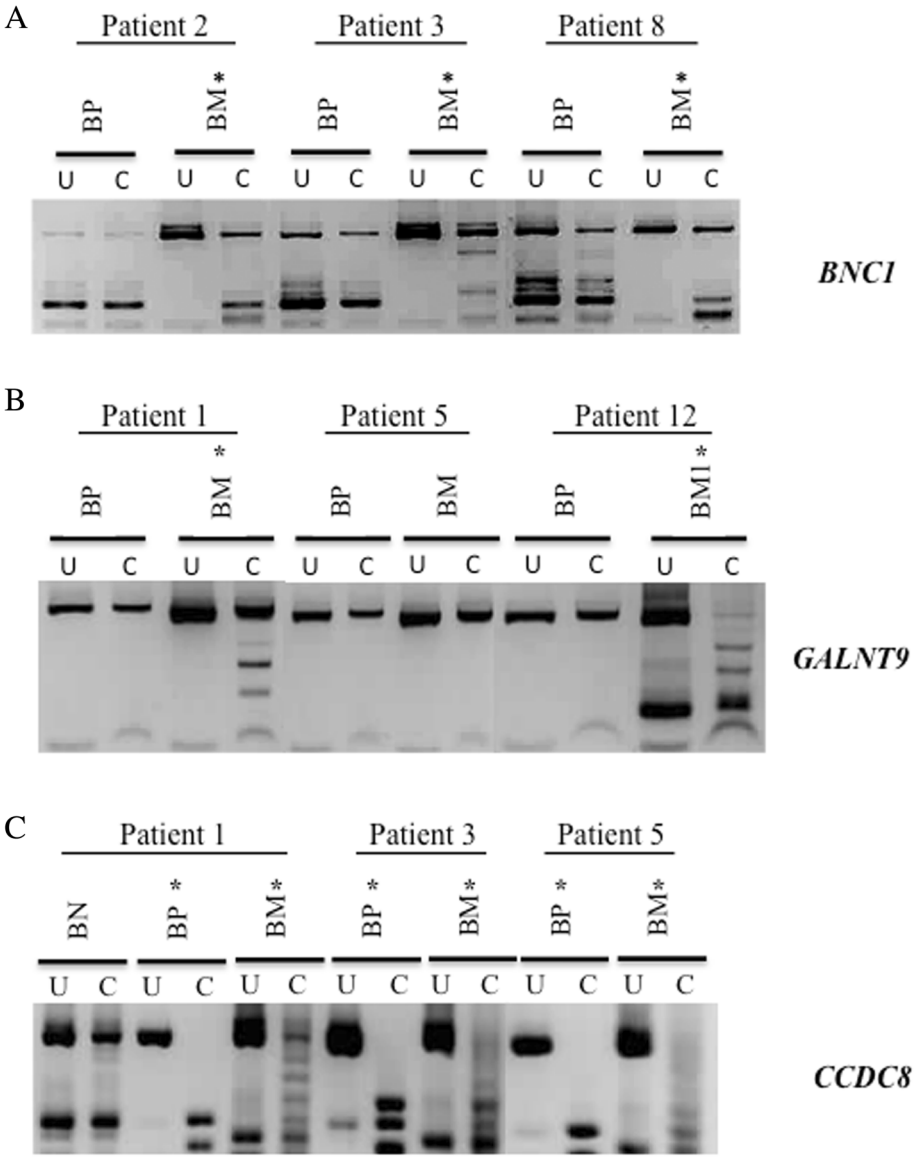


Fig. 2 Methylation status of *GALNT9*, *CCDC8* and *BNC1* in metastatic brain tumours and their corresponding originating primary breast tumours from individual patients. CoBRA was used to determine methylation status; small, digested PCR products in the Bstu1 cut (C) lane compared to the undigested (U) lane indicates promoter methylation in a sample. **a** *GALNT9*, **b** *CCDC8* and **c** *BNC1* were frequently methylated (*) in metastatic brain tumours (BM). However, *GALNT9* and *BNC1* were not commonly methylated in the originating breast primary (BP) tumours (**a**, **c**). *CCDC8* promoter was methylated in both the originating primary tumours (BP) and the associated brain metastases (BM) from individual patients (**b**). Of eight matched pairs analysed, *BNC1* was methylated in all metastatic brain tumours whereas it was methylated in only one of the corresponding primary tumours (for example, see patients 2, 3 and 8). Of six matched pairs analysed, *GALNT9* was methylated in three metastatic brain tumours (see patients 1 and 12), whereas it was not methylated in any of the corresponding primary tumours. Of 11 matched pairs analysed, *CCDC8* was methylated in ten metastatic tumours and all corresponding primary tumours (for example, see patients 1, 3 and 5). However, it was not methylated in normal tissue (BN) adjacent to the primary breast tumour (see patient 1). (*BP* breast primary tumour, *BM* metastatic brain tumour, *BN* adjacent normal breast tissue, *U* uncut/control sample, *C* cut by methylation specific restriction enzyme, *methylated samples)

methylation in primary breast tumour and resulting brain metastasis was confirmed by sequencing individual alleles for pairs of tumours from two patients (patient 11 and 15 (BM11, Primary BT 11 and BM15, Primary BT 15)). Both primary tumour DNA and BM DNA were found to have MIs above 73 % (Additional file 9: Figure S5a).

These results suggest that *BNC1* and *GALNT9* promoter methylation occurs at a late stage in the evolution of metastatic brain tumours, possibly after they have metastasised to the brain. Alternatively, methylation of these genes may occur in a small subset of cells within the primary tumour (below the detection threshold of this assay), and these cells are enriched in the metastatic tumour. In contrast, *CCDC8* promoter methylation is detectable in most primary tumours that metastasise to the brain, suggesting that it may play an important role in the early stages of primary tumour metastasis.

Loss of *GALNT9*, *CCDC8* or *BNC1* expression increases metastatic potential

We have identified that *CCDC8* is dysregulated at an early point of BBM, and its promoter methylation is detectable in the primary tumours that proceed to metastases. *GALNT9* and *BNC1* methylation is uncommon in primary breast tumours and is often not detectable in the tumours that metastasise. These differences suggest that loss of these genes confers metastatic potential though alternative mechanisms. However, loss of *BNC1* or *CCDC8* expression has previously been shown to increase the clonagenic potential of RCC cell lines [31, 32]. Loss of *GALNT9* has yet to be directly linked with increased malignancy. We have investigated the effect that loss of these genes has on metastasis-related properties of breast cancer cell lines.

Loss of *GALNT9*, *CCDC8* or *BNC1* expression increases breast cancer cell line cell motility

Forty-eight hours after initial transfection with siRNA oligos against *BNC1*, *CCDC8* or *GALNT9* breast cancer cell lines showed loss of specific gene expression (Additional file 12: Figure S8).

In a wound-healing assay, knockdown of these genes increased migratory potential compared to cell lines transfected with control oligos. The increase in motility of cell lines following knockdown of *BNC1* (Fig. 3a), *CCDC8* (Fig. 3b) or *GALNT9* (Fig. 3c) was statistically significant compared to control cells (scrambled siRNA transfected)

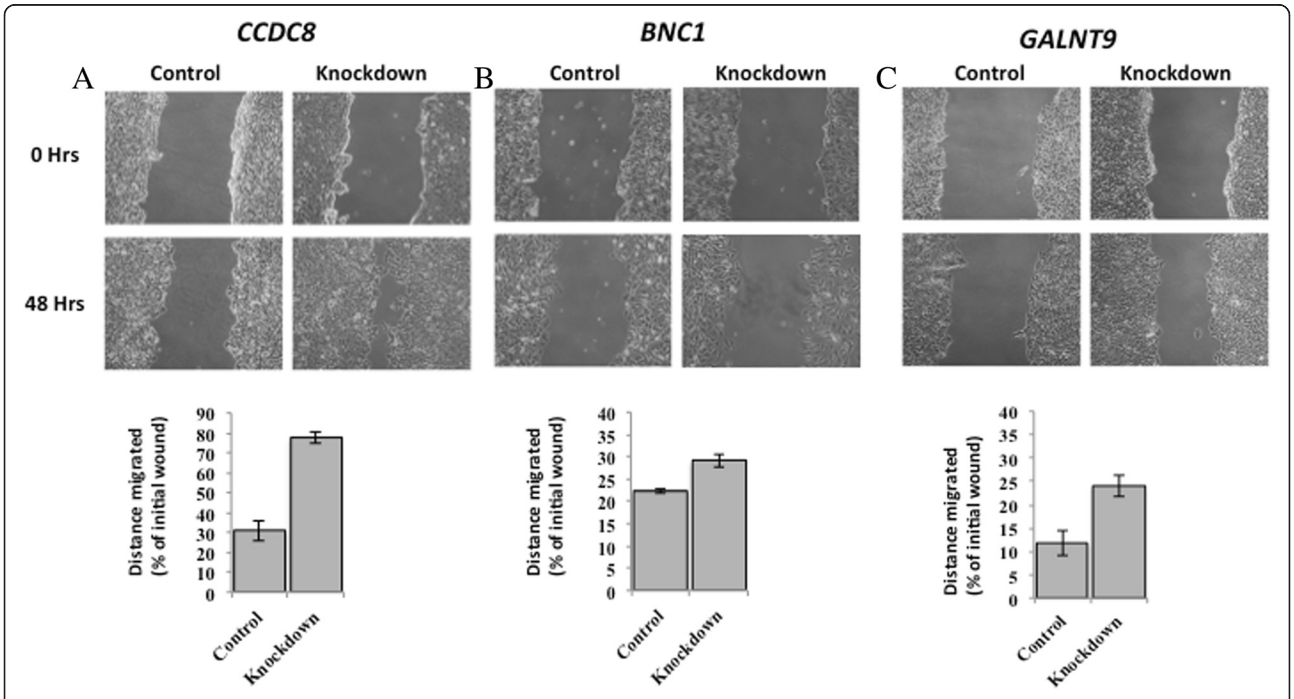


Fig. 3 Loss of *CCDC8*, *BNC1* or *GALNT9* expression increases the migratory potential of breast cancer cell lines. Breast cancer cell lines that expressed *CCDC8*, *BNC1* or *GALNT9* were identified; the expression of these genes was knocked down by siRNA (see Additional file 11: Figure S7 and Additonal file 12: Fig. S8). **a** T47D cell lines transfected with siRNA oligos against *CCDC8*; **b** MCF7 cell lines transfected with siRNA oligos against *BNC1* or **c** MDA-MB231 cell lines transfected with siRNA against *GALNT9* exhibited more migratory potential compared to respective cell lines transfected with control siRNA oligos. Following siRNA transfection, confluent cells were incubated in serum-free media and an artificial wound was scratched through them (0 h). Forty eight hours later the distance migrated was calculated by subtracting the value of non-migrated distance from the initial would. The distance migrated (in percentage) by respective cell lines knocked down with siRNA against *CCDC8*, *BNC1* or *GALNT9* in compared to the respective cell lines transfected with control siRNA oligos was statistically significant ($p = 0.001$, 0.011 and 0.027, respectively)

(BNC1, $p = 0.011$; *CCDC8*, $p = 0.001$; *GALNT9*, $p = 0.027$). All experiments were repeated in triplicate.

Reduced expression of *GALNT9*, *CCDC8* or *BNC1* increases invasive potential

GALNT9, *BNC1* and *CCDC8* were knocked down in breast cancer cell lines by siRNA and applied to matrigel-coated invasion chambers as described in the methods. Forty-eight hours later, cells that had 'invaded' were isolated and quantified.

Following knockdown of *GALNT9*, 35 % more cells invaded ($p = 0.025$) compared to cell transfected with the control scrambled siRNA (Fig. 4a). Following knockdown of *CCDC8*, 27 % more cells invaded, ($p = 0.021$) (Fig. 4b). The number of breast cancer cell that invaded following *BNC1* knocked down was increased by 40 % ($p = 0.006$) (Fig. 4c).

Increased motility and invasive potential following reduction of expression of these genes suggests that these candidates may be involved in the regulation of normal cellular physiology and that loss of their expression may contribute the metastatic process.

Reduced expression of *GALNT9* or *CCDC8* is significantly associated with poor relapse-free survival

The clinical significance of the expression of *BNC1*, *CCDC8* and *GALNT9* was analysed using publically available GEO expression profiles using the prognoscan database [33]. Prognoscan partitions a patient population into high-expressor and low-expressor group for each gene by choosing a threshold that maximises the statistical significance of difference in outcome. It corrects for multiple testing using the method of Miller and Siegmund [34]. In two independent datasets, low *CCDC8* expression was significantly associated with poor relapse free survival (GSE12276, $p = 0.001$; GSE1456-GPL97, $p = 0.004$) (Fig. 5a), and in one data set, low *GALNT9* expression was associated with poor relapse free survival, (GSE1379, $p = 0.0029$) (Fig. 5b). There was no evidence in any of the datasets analysed that low *BNC1* expression correlated with poor relapse free survival or any other clinical indicator.

Discussion

Given the extremely poor clinical outcome following a diagnosis of BBM [9], it is imperative that the underlying molecular biology that drives tumour evolution to the colonization of the brain is elucidated.

To date, some progress has been made to identify prognostic markers for breast cancer metastasis by gene expression profiling [35]. However, prediction of site-specific metastasis remains poor [36].

The importance of gene dysregulation by promoter methylation as a mechanism of tumour evolution is now well established [37]. Indeed, genome wide methylation analysis of many hundreds of primary breast tumours has allowed the definition of specific sub-categories of breast tumours [12, 13], and our increasing understanding of the molecular basis of these subtypes has improved our ability to predict early metastatic recurrence [14, 38]. However, late recurrence, a common feature of BBM has proven difficult to predict [39].

We have carried out a broad candidate approach to identify genes that are dysregulated in BBM (Additional file 1: Figure S1). This analysis has identified three genes (*BNC1*, *CCDC8* and *GALNT9*) that are differentially methylated in primary breast tumours and BBM.

We predicted that our analysis of unrelated primary breast tumours and BBM would identify two different classes of genes that contribute to BBM, epigenetic silencing of BBM associated genes would either occur as (i) *early events* in tumour evolution that may be involved in processes such as local invasion and intravasion [40, 41] or these early events may be required for specific distant site metastasis but also contribute to primary tumour development or (ii) *late events* that play no significant role in the initial evolution of the primary tumour but contribute to the development of the secondary brain metastasis, perhaps by improving the capacity of these cells to survive in the foreign microenvironment of the brain.

The existence of early and late events had previously been proposed by Nguyen et al. [42], they classified deregulated genes as either involved in (i) metastasis initiation, detectable in the primary tumour, (ii) metastasis progression genes, important for survival in the circulation or required for extravasation, while occasionally present in the primary tumour, they may also occur once metastasising cells have left the primary site, or (iii) metastasis virulence genes that allow the cancer cells to survive in a foreign tissue environment. These are likely to occur as a consequence of the selection pressure provided by the novel environment the metastasised tumour cells find themselves in. Metastasis progression genes may have different functions in the primary tumour and distant metastasis, for example, *MMP-1* promotes vascular remodelling in primary breast tumours and also contributes to lung extravasation [43]. An example of a known metastasis virulence genes that does not contribute to primary tumour growth is interleukin-11, which promotes breast tumour metastasis to the bone but does not provide any advantage to the primary tumour [44].

Both early and late methylation events will appear similarly in our initial analysis; the genes will be frequently methylated in BBM and infrequently methylated in unrelated primary breast tumours, this is the case for *BNC1*, *CCDC8* and *GALNT9* (Fig. 1). However, a

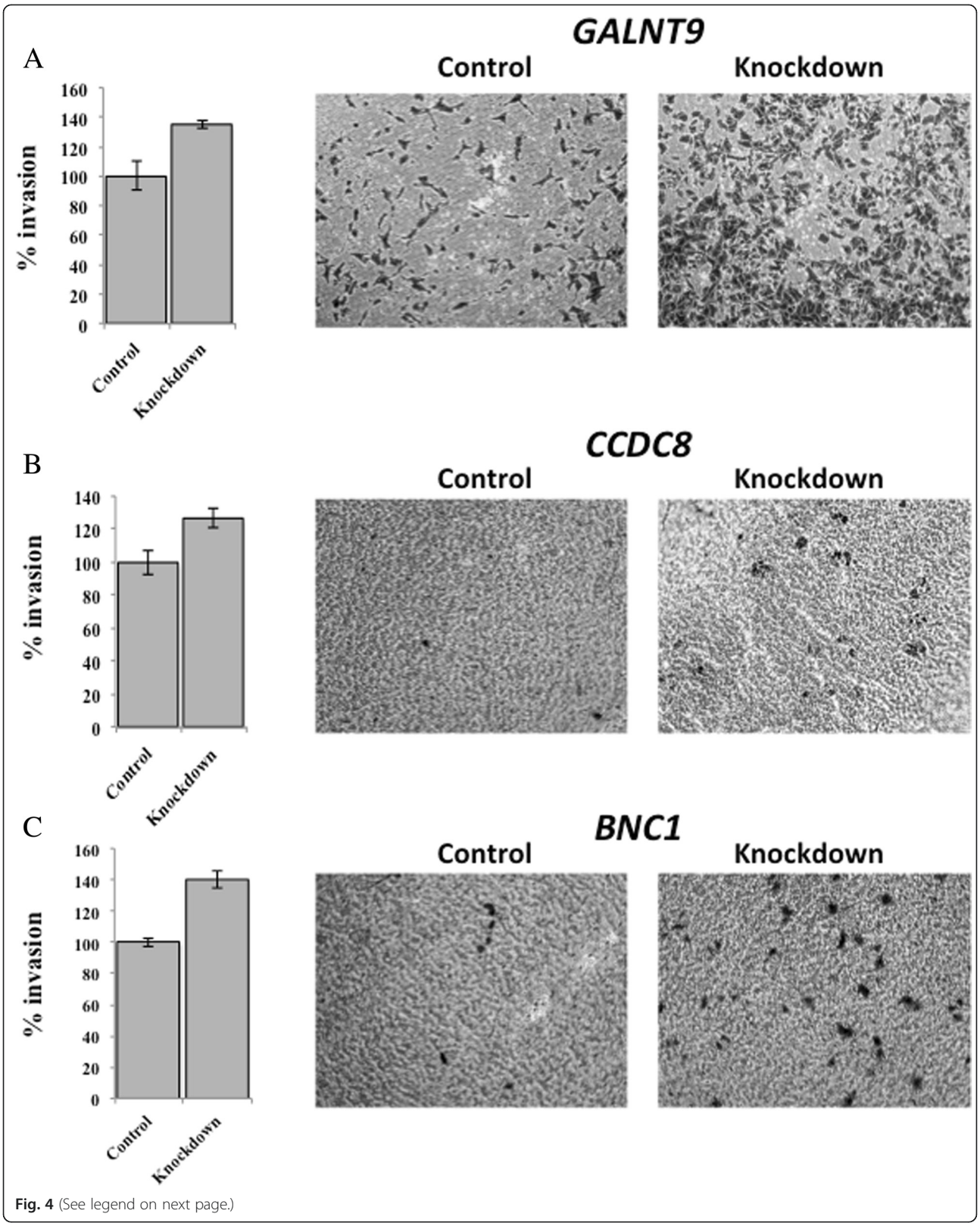


Fig. 4 (See legend on next page.)

(See figure on previous page.)
Fig. 4 Reduced expression of *GALNT9*, *CCDC8* or *BNC1* increases the invasive potential of breast cancer cell lines. Trans-well invasion assays were carried out following the knockdown of of *GALNT9*, *CCDC8* or *BNC1* in breast cancer cell lines. The invasive capacity of these cells was compared with the same cell lines transfected with control siRNA oligos (control). The numbers of cells that had invaded a matrigel-coated micropore membrane was determined colourimetrically 48 h after initial seeding. **a** MDA-MB231 cell lines transfected with siRNA oligos against *GALNT9*, **b** T47D cell lines transfected with siRNA oligos against *CCDC8* and **c** MCF7 cell lines transfected with siRNA oligos against *BNC1* exhibited a statistically significant increase in invasiveness compared to negative control siRNA transfected cells. $p = 0.025$ (*GALNT9*), $p = 0.021$ (*CCDC8*) and $p = 0.001$ (*BNC1*). Invasive potential was calculated as a percentage increase above that observed for the control cells (% invasion)

comparison of primary tumours and BBM from the same patient should reveal if specific gene methylation occurs early or late in the process of tumour evolution. Our analysis of such tumour pairs (Fig. 2) identified that *BNC1* and *GALNT9* are not frequently methylated in any breast tumours, even those that will eventually develop into brain metastases where these genes are methylated. Their methylation appears to be a late event in tumour evolution/metastasis. However, the *CCDC8* promoter was commonly methylated in primary breast tumours that eventually develop brain metastases and as such it can be categorised as an early event in tumour evolution/metastasis.

GALNT9 encodes a member of the UDP-*N*-acetyl-α-D-galactosamine:polypeptide *N*-acetylgalactosaminyl-transferase family of enzymes that catalyze the first step of O-glycosylation; GALNAC-T9. GLANT9 is expressed most abundantly in the brain and other CNS tissues. It is also expressed, at lower levels, in a number of other tissues including normal breast (GeneCards) [45].

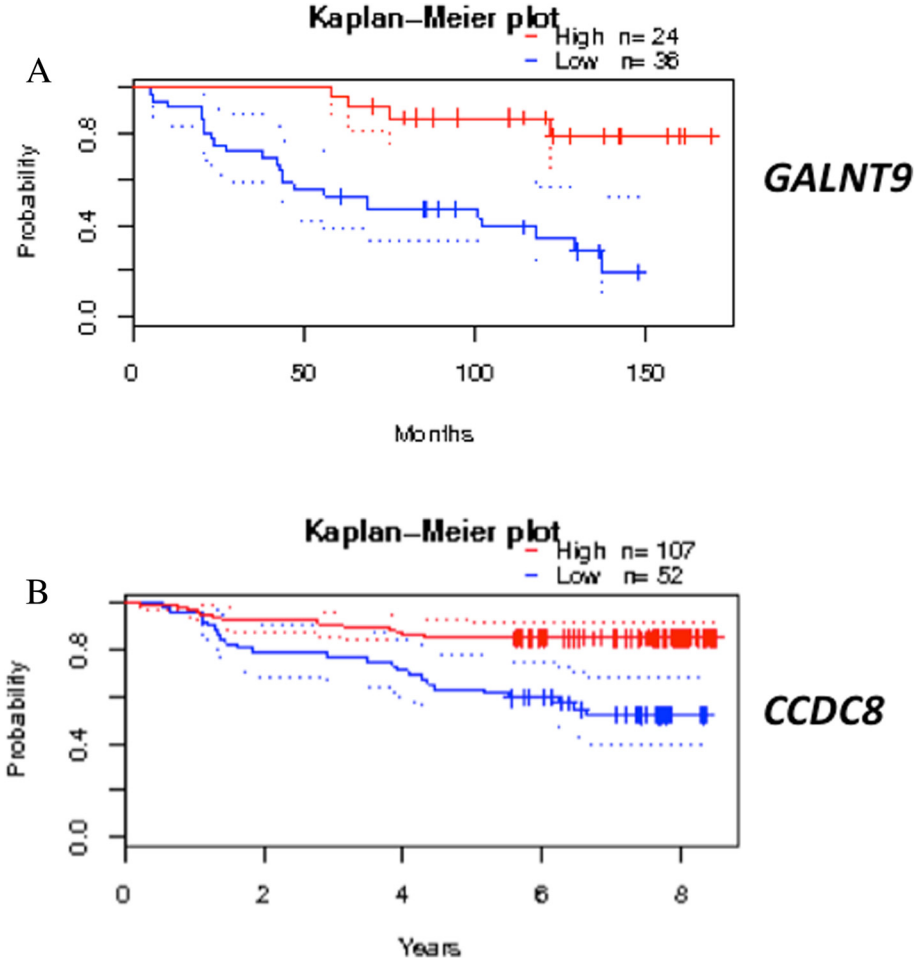


Fig. 5 Loss of expression of *CCDC8* and *GALNT9* correlates with relapse-free patient survival. Kaplan–Meier analysis of multiple gene expression studies via the prognoscan database revealed that in two separate studies low expression of **a** *CCDC8* ($p = 0.001$) and **b** *GALNT9* ($p = 0.003$) was associated with poor relapse free survival

The GALNAC-T proteins initiate mucin type O-linked glycosylation in the golgi apparatus by the covalent linkage of an α-*N*-acetylgalactosamine (GalNAc) to Ser and Thr residues [46]. O-Glycans play an important role in cell adhesion and cell-cell communication, and dysregulated glycosylation is a common characteristic of tumour cells [47]. Mucin 1 (MUC1), in particular, has been identified as a highly O-glycosylated transmembrane protein that is dysregulated at the expression and posttranslational level in multiple tumour types [47]. MUC1 is commonly overexpressed but under-glycosylated in primary breast tumours [48, 49], and the expression of under-glycosylated MUC1 is associated with high tumour grade, metastatic potential and invasiveness of breast tumours [50–52]. Loss of GALNT9 expression in neuroblastoma has been linked to a highly malignant phenotype and associated with poor overall and disease free survival [53]. GALNT9 is a member of a sub family (with GALNT8, 18 and 19) that differ significantly in sequence from other GALNAC-T members [54] and as such does not have catalytic activity towards classic MUC1 variants derivatives [55, 56]. This suggests that GALNT9 glycosylates a specific group of substrates indicating a subtle regulation of transmembrane protein function. Our findings of *GALNT9* promoter methylation, and associated loss of expression, in BBM, but not in primary breast tumours suggest that this change in transmembrane protein function may be a common occurrence in the later stages of breast tumour brain metastasis, and perhaps relates to cell-cell interaction that the tumour cells must undergo before acquiring a suitable niche to proliferate within the novel microenvironment of the brain.

This is the first time that *GLANT9* has been shown to be dysregulated in cancer by promoter methylation. However, conserved mutations have been identified in approximately 2 % of microsatellite instable colorectal cancers [57] and *GALNT9* is also mutated, infrequently (<1 %), in astrocytoma [58] and lung tumours [59, 60] and infrequently lost through CNV in breast tumours [12, 60].

Basonuclin 1 (BNC1) is a zinc finger transcription factor that interacts with the promoters of both RNA polymerases I and II [61]. BNC1 target genes have been implicated in a broad range of functions including chromatin structure, transcription/DNA-binding, adhesion, signal transduction and intracellular transport [61–63]. It is expressed in a broad range of tissue types (GeneCards) [45].

BNC1 has previously been shown to be silenced by promoter methylation in lung [64], renal [31], pancreatic [65], prostate [66] and leukemic cancers [67]. In vitro assays have shown that loss of BNC1 expression is associated with an increased malignant phenotype [31]. Consistent with this study, analysis of HumanMethylation 27 and 450 K array data from *The Cancer Genome Atlas* indicates that *BNC1* Promoter methylation is an infrequent event in primary breast tumours [12]. However, frequent *BNC1* promoter methylation (>60 %) in a small cohort of breast tumours has previously been reported [64].

The expression of BNC1 is induced by transforming growth factor-β1 signalling and, in turn, it acts as a transcription factor for a number of modulators of epithelial dedifferentiation during EMT [68]. Moreover, loss of BNC1 expression results in a reduced EMT phenotype. These findings suggest that the expression of BNC1 would enhance the process of metastasis via EMT. Our findings are consistent with this; we find that *BNC1* is infrequently methylated in primary breast tumours (17 %) and frequently methylated and silenced in BBMs (73 %). Moreover, we have shown that *BNC1* promoter methylation is a late event in tumour evolution, only occurring in the brain metastasis of a BBM patient and not in the associated primary tumour. It is plausible that BNC1 expression is commonly required for EMT to occur during metastasis and, once these cells have metastasised to the brain, loss of *BNC1* expression contributes to mesenchymal to epithelial transition (MET).

An in vitro screen that consisted of multiple rounds of breast cancer cell line injection into nude mice and reculturing of the resulting brain metastases showed that *BNC1* was among a large number of genes overexpressed in mouse brain metastases [69]. This apparent difference to our findings may be as a consequence of the model used. Alternatively, it may represent important differences in the process of aggressive early metastasis (as cell line injection models represent) and slower metastatic evolution, where tumour cells proceed through a phase of latency or micrometastasis. Many of the brain metastases in our study were identified several years after initial breast cancer diagnosis (Paired primary and BBM samples were excised between 2 and 10 years apart).

CCDC8 encodes a coiled-coil domain containing protein (CCDC8) that is one of three proteins that are mutated in patients with 3 M syndrome [70], an autosomal recessive disorder characterised by short stature, skeletal abnormalities, reduced male hormone and blood vessel bulges [71–73]. *CCDC8* is mutated in ~5 % of 3 M cases, the other genes, *CUL7* and *OBSL1* are mutated in ~65 % and ~30 % of cases, respectively [70, 74]. These three proteins form a complex (the 3 M complex) and loss of expression of any one protein disrupts microtubule dynamics resulting in dysregulated mitosis, cytokinesis, associated genomic instability and aneuploidy [75]. Moreover, it was shown that loss of any 3 M complex protein significantly altered the interphase microtubule network [75]. The core 3 M-protein complex interacts with CUL9, which has been proposed to mediate the functions of the 3 M complex via the ubiquitylation and degradation of survivin [76]. The 3 M-complex also interacts with the F

box protein FBXW8, ROC1 and the tumour suppressor p53 [75] suggesting it may contribute to correct cellular physiology through multiple mechanisms.

Despite the broad range and very different known functions that these three proteins have it is interesting to see that, at the level of in vitro assays, reduced expression of any of them increases metastatic potential (Figs. 3, 4).

Conclusions

Our findings indicate that epigenetic dysregulation of *GALNT9*, *CCDC8* or *BNC1* in breast tumours may contribute to metastasis to the brain and possibly other distant organs. *CCDC8* dysregulation occurs early during tumour evolution, in addition to being a potential therapeutic target this early inactivation has the potential to be utilised as a prognostic biomarker. Further analysis will be required including studies to determine if such epigenetic markers can be discerned via non-invasive means such as analysis of circulating tumour material in the patients blood. *GALNT9* and *BNC1* promoter methylation and associated silencing is common in BBM but does not occur frequently in the originating breast tumours suggesting that their dysregulation may not necessarily benefit the primary tumour but are required for successful colonization of the brain. Further studies will be required to determine if these changes are detectable in circulating tumour cells, micrometastases, or only in macroscopic brain metastases. Our current understanding of the cellular function of these genes is far from complete. However, what is known about all three suggests that their dysregulation may be more that just a marker for BBM. As such these genes may represent novel therapeutic targets.

Additional files

Additional file 1: Figure S1. Graphical overview of methodologies used and results obtained in this study. (A) A literature review was carried out to identify genes that are methylated in lung, melanoma and renal cancer as these often metastasise to the brain rapidly. If these genes were not known to be frequently methylated in breast tumours (that metastasise to the brain with a longer lag period) they were considered as good candidates. (B) A literature review was carried out to identify genes down regulated in Epithelial to Mesenchmal Transition (EMT). (C) Analysis of genome-wide methylation data from *The Cancer Genome Atlas* identified 4 genes frequently methylated in Lung tumours and infrequently methylated in breast tumours with no evidence of distant metastasis. Genes from these candidate lists were screened for methylation in breast to brain metastases (BBM), those that were frequently methylated were then screened for methylation in non-metastatic primary breast tumours. Of the 82 genes analysed *BNC1*, *CCDC8* and *GALNT9* were frequently methylated in BBM and infrequently methylated in non-metastatic primary breast tumours, suggesting a role in the evolution of metastatic tumours.

Additional file 2: Table S1. TCGA tumour barcodes: Unique barcode of Breast Invasive Carcinoma (BRCA) and Lung Adenocarcinoma (LUAD) tumours from *The Cancer Genome Atlas* (TCGA) downloaded for bioinformatic analysis to screen for candidate genes, which may contribute to breast to brain metastases (BBM).

Additional file 3: Table S2. Molecular characteristics and other clinical information relating to primary breast tumours analyzed in this study. (A) Primary breast tumours that have metastasised to the brain. These primary tumours have associated metastases analysed in this study. (B) Primary breast tumours with no evidence of metastasis to the brain. These patients have no evidence of developing brain metastases, see Methods: Patients and samples for further details (N: Negative, P: Positive; 0: Negative, 1: Positive).

Additional file 4: Table S3. Primers used in CoBRA and Reverse-Transcription (RT) PCR. CoBRA (Combined Bisulphite and Restriction Analysis) primers were designed to amplify promoter regions of 82 genes In addition, RT primers were designed to amplify transcripts of *BNC1*, *CCDC8* and *GALNT9* to investigate their expression in breast cancer cell lines and breast to brain metastases. F: Forward primer, IF: Internal Forward primer, IR: Internal Reverse primer, R: Reverse primer.

Additional file 5: Figure S2. The promoter region/CpG islands of *BNC1*, *CCDC8* and *GALNT9*. The region amplified for CoBRA analysis is found between the Internal Forward primer and the Reverse primer. CpG dinucleotides are highlighted in bold. CpG dinucleotides analysed by cloning and sequencing of individual alleles are numbered. An arrow indicates the transcription start site.

Additional file 6: Table S4. Genes analysed for their methylation status in breast to brain metastases (BBM) and their function. Methylation status of CpG island promoter region of 82 genes (4 genes from our bioinformatic screen and 78 genes from a broad literature review including genes down regulated in Epithelial- Mesenchymal Transition) was interrogated using Combined Bisulphite and Restriction Analysis (CoBRA) in BBM (n=15). 21 genes were frequently methylated in BBM (light grey background) of which, 3 genes (*CCDC8*, *BNC1* and *GALNT9*) (dark grey background) were infrequently methylated in an independent cohort of primary tumours (n=15). These three genes were further analysed in 20 more primary breast samples (n=30 in total) and 15 more BBM (n=30).

Additional file 7: Figure S3. Methylation analysis of *BNC1*, *CCDC8* and *GALNT9* in Breast to brain metastases. Up to 31 brain metastases (BM) were analysed by CoBRA, small, digested PCR products in the Bstu1 cut (C) lane compared to the undigested (U) lane indicates promoter methylation in a sample (SAM DNA: genomic DNA treated with *S*-Adenosyl methionine and DNA methyltransferase as a positive control).

Additional file 8: Figure S4. Methylation status of *GALNT9*, *CCDC8* and *BNC1* in metastatic brain tumours from primary breast tumours and a cohort of unrelated primary breast tumours. (A) *GALNT9* was frequently methylated in metastatic brain tumours (55 %) and (B) was NOT methylated in any of the primary tumours; (C) *CCDC8* was frequently methylated in metastatic brain tumours (73 %) and (D) infrequently methylated in primary breast tumours (40 %). (E) *BNC1* is frequently methylated in metastatic brain tumours 68 % and (F) infrequently methylated in a cohort of unrelated primary breast tumours (17 %). We have determined that a significant proportion of the promoters within the tumour sample are methylated if there are clearly observed digest products following restriction analysis BM: Brain metastases, BP: Primary breast tumours, U: Uncut/control sample, C: cut by restriction enzyme, *: methylated samples).

Additional file 9: Figure S5. Bisulphite sequencing of individual alleles form tumours. Tumours were analysed by cloning and sequencing bisulphite-PCR products to determine the extent of methylation within the region analysed by CoBRA. 10 clones/alleles were sequenced for each tumour and the methylation index (MI) for each tumour determined. (A) Tumours that were determined to have significant *CCDC8* promoter methylation by CoBRA (BM11, BM15, BM12 and BM14) had methylation indices ranging from 66 %-90 %. The corresponding primary breast tumours for BM11 and BM15 were also analyses these both had correspondingly high MIs (82 % and 74 % respectively). Tumours that had no evidence of *CCDC8* promoter region methylation by CoBRA analysis (BM16, BM23) had low MIs (8 % and 6 % respectively). (B) Tumours that were determined to have significant *BNC1* promoter methylation by CoBRA (BM13, BM14, BM15 and BM27) had methylation indices ranging from 60 %-86 %. Tumours that had no evidence of *BNC1* promoter region methylation by CoBRA analysis

(BM11, BM23) had low MIs (0 % and 36 % respectively). (C) Tumours that were determined to have significant *GALNT9* promoter methylation by CoBRA (BM12, BM20, BM27 and BM28) had methylation indices ranging from 78 %-91 %. Tumour BM23 that had no evidence of *GALNT9* promoter region methylation by CoBRA analysis had a low MI (25 %). Each circle represents a CpG island, those shaded black are methylated. MI is defended as the total number of methylated CpG dinucleotides given as a percentage of all CpGs analysed.

Additional file 10: Figure S6. Expression levels of *BNC1. CCDC8* and *GALNT9* in all tumours analysed. The expression level of each gene was quantified in relation to the expression of β-actin. Below each bar is the methylation status of each CpG island as determined by CoBRA and sequencing of individual alleles (MI) (BM: Brain Metastasis, MI: Methylation index, M: Methylated, U: Unmethylated, -: analysis was not done).

Additional file 11: Figure S7. Global demethylation resulted in the re-expression of *GALNT9*, *CCDC8* and *BNC1* in breast cancer cell lines. Reverse transcription PCR (RT-PCR) showed that treatment of breast cancer cell lines with 5- 2-deoxycytidine (5-AZA-dC), an inhibitor of DNA methyltransferase enzymes, resulted in re-expression of (A) *GALNT9,* (B) *CCDC8* or (C) *BNC1* in the breast cancer cell line ZR75. For comparison, endogenous expression is shown in (A) MDA-MD231, (B) T47D and (C) MCF7, these, expressing, cell lines were used in our *in vitro* knock down experiments.

Additional file 12: Figure S8. Knockdown of *GALNT9*, *CCDC8,* and *BNC1* in breast cancer cell lines is confirmed by Reverse transcription (RT) PCR and western blot. (A) RT-PCR of *GALNT9*, *CCDC8,* and *BNC1* transcripts in breast cancer cell lines (MDA-MD231, T47D and MCF7 respectively) following siRNA knockdown compared to transfection with a control siRNA and (B) western blot of GALNT9, CCDC8, and BNC1 proteins to confirm their knockdown in each respective cell line. 70 μg of protein was loaded in each lane. Equal loading was confirmed by staining total protein with India ink.

Abbreviations

5-AZA-dC: 5-AZA-2′-deoxycytidine; BBM: Breast to brain metastasis; BM: Brain metastasis; BN: Normal tissue adjacent to breast tumour; BP: Breast primary tumour; CoBRA: Combined bisulfite and restriction analysis; EMT: Epithelial to mesenchymal Transition; ER: Estrogen receptor; FFPE: Formalin fixed paraffin embedded; HER2: Human epidermal growth factor receptor 2 (ERBB2); MET: Mesenchymal to epithelial transition; MI: Methylation index; PR: Progesterone receptor; RCC: Renal cell carcinoma; SAM: S-adenosyl methionine; TCGA: The Cancer Genome Atlas.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

RPP carried out preliminary and final CoBRA methylation analysis, expression analysis and all cell culture/functional work; PC carried out cloning and sequencing of alleles; D.S.H. carried out statistical analysis and prognoscan analysis; BKJ, DP, LWE, NH, ON, CLD, AIA, HK, ISS, VCB, PC and TS carried out preliminary Candidate CoBRA screening; TPD, KMA, CD, ARB, MDJ and IB provided tumour material and clinical information; FL, JLD, TJW and MRM designed and supervised experiments and analysis of data; MRM directed the research; RPP and MRM drafted the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We would like to thank the Sidney Driscol Neuroscience Foundation for supporting this research by funding the BTNW research tissue bank.

Author details

[1]Brain Tumour Research Centre, University of Wolverhampton, Wolverhampton, UK. [2]School of Biology, Chemistry and Forensic Sciences, University of Wolverhampton, Wolverhampton, UK. [3]Centre for Rare Diseases and Personalised Medicine, School of Clinical and Experimental Medicine, University of Birmingham, Birmingham, UK. [4]Department of Neurosciences, Lancashire Teaching Hospitals NHS Foundation Trust, Royal Preston Hospital, Fulwood, Preston, UK. [5]The Walton Centre NHS Foundation Trust, Lower Lane, Liverpool, UK. [6]Department of Genetics, Institute Curie, Paris, France.

Received: 16 December 2014 Accepted: 11 May 2015
Published online: 27 May 2015

References

1. Gavrilovic IT, Posner JB. Brain metastases: Epidemiology and pathophysiology. J Neuro-Oncol. 2005;75(1):5–14.
2. Soffietti R, Ruda R, Mutani R. Management of brain metastases. J Neurol. 2002;249(10):1357–69.
3. Gori S et al. Central nervous system metastases in HER-2 positive metastatic breast cancer patients treated with trastuzumab: Incidence, survival, and risk factors. Oncologist. 2007;12(7):766–73.
4. Kennecke HF et al. Late risk of relapse and mortality among postmenopausal women with estrogen responsive early breast cancer after 5 years of tamoxifen. Ann Oncol. 2007;18(1):45–51.
5. Tosoni A, Franceschi E, Brandes AA. Chemotherapy in breast cancer patients with brain metastases: have new chemotherapic agents changed the clinical outcome? Crit Rev Oncol Hematol. 2008;68(3):212–21.
6. Weil RJ et al. Breast cancer metastasis to the central nervous system. Am J Pathol. 2005;167(4):913–20.
7. Soffietti R, Ruda R, Trevisan E. Brain metastases: Current management and new developments. Curr Opin Oncol. 2008;20(6):676–84.
8. Jenkinson MD et al. Management of cerebral metastasis: Evidence-based approach for surgery, stereotactic radiosurgery and radiotherapy. Eur J Cancer. 2011;47(5):649–55.
9. Sperduto PW et al. Diagnosis-specific prognostic factors, indexes, and treatment outcomes for patients with newly diagnosed brain metastases: a multi-institutional analysis of 4,259 patients. Int J Radiat Oncol Biol Phys. 2010;77(3):655–61.
10. SUGARBAKER ED. The organ selectivity of experimentally induced metastases in rats. Cancer. 1952;5(3):606–12.
11. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. Nature. 2009;458(7239):719–24.
12. TCGA. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490(7418):61–70.
13. Curtis C et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature. 2012;486(7403):346–52.
14. Metzger FO, Ignatiadis M, Sotiriou C. Genomic Grade Index: an important tool for assessing breast cancer tumor grade and prognosis. Crit Rev Oncol Hematol. 2011;77(1):20–9.
15. Sanna G et al. Brain metastases in patients with advanced breast cancer. Anticancer Res. 2007;27(4C):2865–9.
16. van't Veer LJ et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature. 2002;415(6871):530–6.
17. Ryberg M et al. Predictors of central nervous system metastasis in patients with metastatic breast cancer. A competing risk analysis of 579 patients treated with epirubicin-based chemotherapy. Breast Cancer Res Treat. 2005;91(3):217–25.
18. Zhang XH et al. Metastasis dormancy in estrogen receptor-positive breast cancer. Clin Cancer Res. 2013;19(23):6389–97.
19. Aguirre-Ghiso JA. Models, mechanisms and clinical evidence for cancer dormancy. Nat Rev Cancer. 2007;7(11):834–46.
20. Lim E, Metzger-Filho O, Winer EP. The natural history of hormone receptor-positive breast cancer. Oncology (Williston Park). 2012;26(8):688–94. 696.
21. Pantel K, Brakenhoff RH, Brandt B. Detection, clinical relevance and specific biological properties of disseminating tumour cells. Nat Rev Cancer. 2008;8(5):329–40.
22. Klein CA. The systemic progression of human cancer: a focus on the individual disseminated cancer cell–the unit of selection. Adv Cancer Res. 2003;89:35–67.
23. Wang Y, Shang Y. Epigenetic control of epithelial-to-mesenchymal transition and cancer metastasis. Exp Cell Res. 2013;319(2):160–9.
24. Joyce JA, Pollard JW. Microenvironmental regulation of metastasis. Nat Rev Cancer. 2009;9(4):239–52.
25. Mundy GR. Metastasis to bone: Causes, consequences and therapeutic opportunities. Nat Rev Cancer. 2002;2(8):584–93.
26. Park BK et al. NF-kappaB in breast cancer cells promotes osteolytic bone metastasis by inducing osteoclastogenesis via GM-CSF. Nat Med. 2007;13(1):62–9.

27. Hill VK et al. Identification of 5 novel genes methylated in breast and other epithelial cancers. Mol Cancer. 2010;9:51.
28. Clark SJ et al. High sensitivity mapping of methylated cytocines. Nucleic Acid Research. 1994;22(15):2990–7.
29. Eichler AF et al. The biology of brain metastases-translation to new therapies. Nat Rev Clin Oncol. 2011;8(6):344–56.
30. Patra SK, Bettuzzi S. Epigenetic DNA-(cytosine-5-carbon) modifications: 5-aza-2′-deoxycytidine and DNA-demethylation. Biochemistry (Mosc). 2009;74(6):613–9.
31. Morris MR et al. Identification of candidate tumour suppressor genes frequently methylated in renal cell carcinoma. Oncogene. 2010;29(14):2104–17.
32. Morris MR et al. Genome-wide methylation analysis identifies epigenetically inactivated candidate tumour suppressor genes in renal cell carcinoma. Oncogene. 2011;30(12):1390–401.
33. Mizuno H et al. PrognoScan: a new database for meta-analysis of the prognostic value of genes. BMC Med Genomics. 2009;2:18.
34. Miller R, Siegmund D. Maximally selected chi square statistics. Biometrics. 1982;38:1011–6.
35. Paik S et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med. 2004;351(27):2817–26.
36. Weigelt B, Peterse JL, van 't Veer LJ. Breast cancer metastasis: markers and models. Nat Rev Cancer. 2005;5(8):591–602.
37. Cock-Rada A, Weitzman JB. The methylation landscape of tumour metastasis. Biol Cell. 2013;105(2):73–90.
38. Stephens PJ et al. The landscape of cancer genes and mutational processes in breast cancer. Nature. 2012;486(7403):400–4.
39. Burstein HJ, Griggs JJ. Deep time: the long and the short of adjuvant endocrine therapy for breast cancer. J Clin Oncol. 2012;30(7):684–6.
40. Gupta GP, Massague J. Cancer metastasis: building a framework. Cell. 2006;127(4):679–95.
41. Chambers AF, Groom AC, MacDonald IC. Dissemination and growth of cancer cells in metastatic sites. Nat Rev Cancer. 2002;2(8):563–72.
42. Nguyen DX, Bos PD, Massague J. Metastasis: from dissemination to organ-specific colonization. Nat Rev Cancer. 2009;9(4):274–84.
43. Minn AJ et al. Genes that mediate breast cancer metastasis to lung. Nature. 2005;436(7050):518–24.
44. Kang Y et al. A multigenic program mediating breast cancer metastasis to bone. Cancer Cell. 2003;3(6):537–49.
45. Safran M et al. GeneCards Version 3: the human gene integrator. Database. 2010. doi:10.1093/database/baq020.
46. Hanisch FG. O-glycosylation of the mucin type. Biol Chem. 2001;382(2):143–9.
47. Brockhausen I. Mucin-type O-glycans in human colon and breast cancer: Glycodynamics and functions. EMBO Rep. 2006;7(6):599–604.
48. Perey L, Hayes DF, Kufe D. Effects of differentiating agents on cell surface expression of the breast carcinoma-associated DF3-P epitope. Cancer Res. 1992;52(22):6365–70.
49. Hayes DF et al. Prediction of prognosis in primary breast cancer by detection of a high molecular weight mucin-like antigen using monoclonal antibodies DF3, F36/22, and CU18: a Cancer and Leukemia Group B study. J Clin Oncol. 1991;9(7):1113–23.
50. van der Vegt B et al. The expression pattern of MUC1 (EMA) is related to tumour characteristics and clinical outcome of invasive ductal breast carcinoma. Histopathology. 2007;51(3):322–35.
51. de Roos MA et al. The expression pattern of MUC1 (EMA) is related to tumour characteristics and clinical outcome in 'pure' ductal carcinoma in situ of the breast. Histopathology. 2007;51(2):227–38.
52. Ghosh SK et al. Expression of underglycosylated MUC1 antigen in cancerous and adjacent normal breast tissues. Clin Breast Cancer. 2013;13(2):109–18.
53. Berois N et al. GALNT9 gene expression is a prognostic marker in neuroblastoma patients. Clin Chem. 2013;59(1):225–33.
54. Raman J et al. UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferases: completion of the family tree. Glycobiology. 2012;22(6):768–77.
55. Li X et al. Characterization of ppGalNAc-T18, a member of the vertebrate-specific Y subfamily of UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferases. Glycobiology. 2012;22(5):602–15.
56. Zhang Y et al. Cloning and characterization of a new human UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase, designated pp-GalNAc-T13, that is specifically expressed in neurons and synthesizes GalNAc alpha-serine/threonine antigen. J Biol Chem. 2003;278(1):573–84.
57. Tuupanen S et al. Identification of 33 candidate oncogenes by screening for base-specific mutations. Br J Cancer. 2014;111(8):1657–62.
58. Brennan CW et al. The somatic genomic landscape of glioblastoma. Cell. 2013;155(2):462–77.
59. Research CGA. N. Comprehensive molecular profiling of lung adenocarcinoma. Nature. 2014;511(7511):543–50.
60. Research CGA, N. Comprehensive genomic characterization of squamous cell lung cancers. Nature. 2012;489(7417):519–25.
61. Zhang S, Wang J, Tseng H. Basonuclin regulates a subset of ribosomal RNA genes in HaCaT cells. PLoS One. 2007;2(9), e902.
62. Wang J et al. Search for basonuclin target genes. Biochem Biophys Res Commun. 2006;348(4):1261–71.
63. Ma J et al. Basonuclin: a novel mammalian maternal-effect gene. Development. 2006;133(10):2053–62.
64. Shames DS et al. A genome-wide screen for promoter methylation in lung cancer identifies novel methylation markers for multiple malignancies. PLoS Med. 2006;3(12), e486.
65. Yi JM et al. Novel methylation biomarker panel for the early detection of pancreatic cancer. Clin Cancer Res. 2013;19(23):6544–55.
66. Devaney JM et al. Identification of novel DNA-methylated genes that correlate with human prostate cancer and high-grade prostatic intraepithelial neoplasia. Prostate Cancer Prostatic Dis. 2013;16(4):292–300.
67. Dunwell TL et al. Epigenetic analysis of childhood acute lymphoblastic leukemia. Epigenetics. 2009;4(3):185–93.
68. Feuerborn A et al. Basonuclin-1 modulates epithelial plasticity and TGF-beta1-induced loss of epithelial cell integrity. Oncogene. 2015;34(9):1185–95.
69. Guo L et al. Selection of brain metastasis-initiating breast cancer cells determined by growth on hard agar. Am J Pathol. 2011;178(5):2357–66.
70. Hanson D et al. Exome sequencing identifies CCDC8 mutations in 3-M syndrome, suggesting that CCDC8 contributes in a pathway with CUL7 and OBSL1 to control human growth. Am J Hum Genet. 2011;89(1):148–53.
71. Huber C et al. Identification of mutations in CUL7 in 3-M syndrome. Nat Genet. 2005;37(10):1119–24.
72. Huber C et al. A large-scale mutation search reveals genetic heterogeneity in 3M syndrome. EJHG. 2009;17(3):395–400.
73. Maksimova N et al. Clinical, molecular and histopathological features of short stature syndrome with novel CUL7 mutation in Yakuts: new population isolate in Asia. J Med Genet. 2007;44(12):772–8.
74. Huber C, Munnich A, Cormier-Daire V. The 3M syndrome. Best Pract Res Clin Endocrinol Metab. 2011;25(1):143–51.
75. Yan J et al. The 3M complex maintains microtubule and genome integrity. Mol Cell. 2014;54(5):791–804.
76. Li Z et al. CUL9 mediates the functions of the 3M complex and ubiquitylates survivin to maintain genome integrity. Mol Cell. 2014;54(5):805–19.
77. Tsukita S, Furuse M. The Structure and Function of Claudins, Cell Adhesion Molecules at Tight Junctions. Ann N Y Acad Sci. 2000;915:129–35.
78. Bragulla HH, Homberger DG. Structure and functions of keratin proteins in simple, stratified, keratinized and cornified epithelia. J Anat. 2009;214(4):516–59.
79. Shimomura Y et al. Mutations in the keratin 85 (KRT85/hHb5) gene underlie pure hair and nail ectodermal dysplasia. J Invest Dermatol. 2010;130(3):892–5.
80. Magin TM, Vijayaraj P, Leube RE. Structural and regulatory functions of keratins. Exp Cell Res. 2007;313(10):2021–32.
81. Chen J et al. miR-127 regulates cell proliferation and senescence by targeting BCL6. PLoS One. 2013;8(11):e80266.
82. Luo H et al. Down-regulated miR-9 and miR-433 in human gastric carcinoma. J Exp Clin Cancer Res. 2009;28:82.
83. Symmans WF. Tissue banking. Breast Cancer Res. 2010;12 Suppl 1:L1.
84. Chen Y et al. Retroviral delivery of homeobox D3 gene induces cerebral angiogenesis in mice. J Cereb Blood Flow Metab. 2004;24(11):1280–7.
85. Pellegrino L et al. miR-23b regulates cytoskeletal remodeling, motility and metastasis by directly targeting multiple transcripts. Nucleic Acids Res. 2013;41(10):5400–12.
86. Jin L et al. Prooncogenic factors miR-23b and miR-27b are regulated by Her2/Neu, EGF, and TNF-alpha in breast cancer. Cancer Res. 2013;73(9):2884–96.
87. Majid S et al. miR-23b represses proto-oncogene Src kinase and functions as methylation-silenced tumor suppressor with diagnostic and prognostic significance in prostate cancer. Cancer Res. 2012;72(24):6435–46.

88. Zaman MS et al. Inhibition of PTEN gene expression by oncogenic miR-23b-3p in renal cancer. PLoS One. 2012;7(11), e50203.
89. Kim YR et al. HOXB13 promotes androgen independent growth of LNCaP prostate cancer cells by the activation of E2F signaling. Mol Cancer. 2010;9:124.
90. Muggerud AA et al. Frequent aberrant DNA methylation of ABCB1, FOXC1, PPP2R2B and PTEN in ductal carcinoma in situ and early invasive breast cancer. BCR. 2010;12(1):R3.
91. Rosen H, Krichevsky A, Polakiewicz RD, Benzakine S, Bar-Shavit Z. Developmental regulation of proenkephalin gene expression in osteoblasts. Mol Endocrinol. 2013;9(11):1621–31.
92. Wagh PK, Peace BE, Waltz SE. Met-Related Receptor Tyrosine Kinase Ron in Tumor Growth and Metastasis. Adv Cancer Res. 2008;100:1–33.
93. Sabine Strehl KG, Qiu ML, Heather G, Lalande M. Characterization of Two Novel Protocadherins (PCDH8 and PCDH9) Localized on Human Chromosome 13 and Mouse Chromosome 14. Genomics. 1998;53:81–9.
94. Niwa H, Burdon T, Chambers I, Smith A. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. Genes Dev. 1998;12:2048–60.
95. Takeda K, Shi W, Tanaka T, Matsumoto M, Yoshida N, Kishimoto T, et al. Targeted disruption of the mouse Stat3 gene leads to early embryonic lethality. Dev Biol. 1997;94:3801–4.
96. Akira S. Roles of STAT3 defined by tissue-specifc gene targeting. Oncogene. 2000;19(21):2607–11.
97. Osler ME, Smith TK, Bader DM. Bves, a member of the Popeye domain-containing gene family. Dev Dyn. 2006;235(3):586–93.
98. Hill VK et al. Genome-wide DNA methylation profiling of CpG islands in breast cancer identifies novel genes associated with tumorigenicity. Cancer Res. 2011;71(8):2988–99.
99. Osanai M et al. Epigenetic silencing of claudin-6 promotes anchorage-independent growth of breast carcinoma cells. Cancer Sci. 2007;98(10):1557–62.
100. Carrio M et al. Homeobox D10 induces phenotypic reversion of breast tumor cells in a three-dimensional culture model. Cancer Res. 2005;65(16):7177–85.
101. Shinya Toba MT, Morichika KTM, et al. Brain-speci¢c expression of a novel human UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase (GalNAc-T9). Biochim Biophys Acta. 2000;1493:264–8.
102. Ai L et al. Inactivation of Wnt inhibitory factor-1 (WIF1) expression by epigenetic silencing is a common event in breast cancer. Carcinogenesis. 2006;27(7):1341–8.
103. Veeck J et al. Prognostic relevance of Wnt-inhibitory factor-1 (WIF1) and Dickkopf-3 (DKK3) promoter methylation in human breast cancer. BMC Cancer. 2009;9:217.